

UNIVERSIDAD COMPLUTENSE DE MADRID
FACULTAD DE CIENCIAS ECONÓMICAS Y EMPRESARIALES



TESIS DOCTORAL

**Aplicaciones de la teoría del control óptimo a la planificación
económica**

MEMORIA PARA OPTAR AL GRADO DE DOCTOR
PRESENTADA POR

José, Borrell Fontelles

DIRECTOR:

Manuel López Cachero

Madrid, 2015

APLICACIONES DE LA TEORIA DEL CONTROL OPTIMO

A LA PLANIFICACION ECONOMICA

Tesis doctoral presentada en -
la Facultad de Ciencias Econó-
micas de la Universidad Complu-
tense de Madrid, bajo la direc-
ción del Profesor Dr. Manuel
Lopez Cachero por

Jose Borrell Fontelles.

Somosaguas, Mayo de 1976.



"Los economistas suelen utilizar las matemáticas como los borrachos las farolas: para apoyarse en ellas más que para iluminarse".

Dicho popular.

INDICE

PAGINA

CAPITULO I

SINTESIS Y PERSPECTIVA

I.1	PRESENTACION	1
I.2	SINTESIS	3
I.3	PERSPECTIVA	9

CAPITULO II

INTERPRETACION ECONOMICA DE LA TEORIA DEL

CONTROL OPTIMO

II.1	INTRODUCCION	18
II.2	OPTIMIZACION Y EQUILIBRIO	21
II.3	SISTEMAS DE VALOR Y PERSPECTIVAS DE ACCION	24
II.4	EL SISTEMA OPTIMO DE VALORACION	29
II.5	LA DESCENTRALIZACION TEMPORAL DE LAS DECISIONES	33
II.6	VALORACION DE LAS RESTRICCIONES	37
II.7	VARIABLES ADJUNTAS Y ANALISIS DE SENSIBILIDAD	40
II.8	CENTROS DE DECISION Y OBJETIVOS MULTIPLES	48
II.9	ALCANCES PRACTICOS	57

CAPITULO III

METODOS DE SOLUCION

III.1	INTRODUCCION	63
III.2	METODOS DIRECTOS Y PROGRAMACION MATEMATICA	67
III.3	METODO PRIMAL DE RESOLUCION: UN ALGORITMO DEL TIPO GRADIENTE REDUCIDO GENERALIZADO	75
3.1	Introducción	75
3.2	Estructura básica de los métodos GRG	77
3.3	Aplicación a problemas discretos de control óptimo. Cálculo del gradiente reducido.	90
3.4	Conclusiones	115

III.4	METODOS DUALES DE RESOLUCION	119
4. 1	Introducción a los métodos duales	119
4. 2	Algoritmo basado en el método dual de los multiplicadores de Hestenes	125
4. 3	Algoritmo dual puro tipo Lasdon-Tamura	141

CAPITULO IV

CONTROL OPTIMO Y PLANIFICACION DINAMICA MULTISECTORIAL CON CENTRO DE DECISION UNICO

IV.1	PRESENTACION	150
IV.2	EL MODELO LINEAL-CUADRATICO	153
2. 1	Reglas de decisión lineales y "feed-back" control	153
2. 2	Características del sistema lineal	166
2. 3	Características de las funciones objetivo cuadráticas	175
IV.3	MODELO DE PLANIFICACION DINAMICA MULTISECTORIAL NO LINEAL	182
3. 1	Planteamiento del modelo	182
3. 2	Formulación del problema de control	207
3. 3	Resolución por el método dual de los multiplicadores	213
IV.4	LOS LIMITES DE UNA SOLUCION GLOBAL	219

CAPITULO V

DESCENTRALIZACION ESTRUCTURAL POR DESCOMPOSICION JERAR- QUICA DEL PROBLEMA DE CONTROL

V.1	INTRODUCCION. LOS SISTEMAS MULTINIVELES DE CONTROL	223
V.2	ALGORITMOS JERARQUICOS DE CONTROL OPTIMO	229
V.3	PLANTEAMIENTO DESCENTRALIZADO DEL MODELO DINAMICO DE PLANIFICACION	242
V.4	POSIBILIDADES, LIMITES E INTERES DE LA APLICACION DEL METODO	262

APENDICE I

A1.1

NOMENCLATURA DE UN PROBLEMA CONTINUO DE CONTROL
OPTIMO Y FORMULACION GENERAL DEL PRINCIPIO DEL
MAXIMO DE PONTRYAGIN.

APENDICE II

A2.1

NOMENCLATURA DE UN PROBLEMA DISCRETO DE CONTROL
OPTIMO Y FORMULACION DEL PRINCIPIO DEL MAXIMO
DISCRETO.

APENDICE III

A3.1

METODOS DE OPTIMIZACION NO LINEAL.

APENDICE IV

A4.1

EL PRINCIPIO DE CERTEZA-EQUIVALENCIA EN EL CON-
TROL OPTIMO DE SISTEMAS LINEAL CUADRATICOS ESTO-
CASTICOS.

INDICE DE FIGURAS

<u>CAPITULO</u>	<u>FIGURA</u>		<u>PAGINA</u>
III	FIII.1	Esquema general de los métodos GRG.	82
III	FIII.2	Estructura de la matriz de base de un problema de control óptimo.	103
III	FIII.3	Organigrama del cálculo del gradiente reducido en un problema de control óptimo.	106-114
III	FIII.4	Organigrama del método de Héstenes.	127-128
III	FIII.5	Organigrama del método dual de los multiplicadores.	133-135
III	FIII.6	Estructura del Jacobiano de las ecuaciones de estado cuando no hay efectos retardados.	138
III	FIII.7	Estructura del Jacobiano de las ecuaciones de estado, cuando existen efectos retardados en estados y controles.	139
III	FIII.8	Flujo de información primal dual en la descomposición dinámica, por un método Lasdon-Tamura, de un problema de control óptimo.	148
IV	FIV.1	Esquema del desglose en tres sistemas de relaciones de las ecuaciones dinámicas de input-output.	174
IV	FIV.2	Productividades marginales del factor trabajo para dos sistemas de dotaciones de capital, en cuatro sectores de actividad del modelo de Kendrick y Taylor.	195
IV	FIV.3	Representación gráfica de los datos de la figura FIV.2.	196
V	FV.1	Esquema de la estructura de un sistema compuesto por subsistemas interrelacionados.	225
V	FV.2	Flujos de entrada y salida de un subsistema.	225

<u>CAPITULO</u>	<u>FIGURA</u>		<u>PAGINA</u>
V	FV.3	Esquema del control de un proceso global por coordinación jerárquica a tres niveles	226
V	FV.4	Esquema del proceso a dos niveles de descomposición jerárquica estructural del modelo de planificación.	253
V	FV.5	Esquema del proceso a tres niveles de descomposición jerárquica estructural y temporal del modelo de planificación.	254
V	FV.6	Matriz de las restricciones de interconexión del modelo de planificación.	250
V	FV.7	Representación del conjunto C de un problema convexo.	274
V	FV.8	Ilustración del teorema del hiperplano minimizante	274
V	FV.9	Caso de un problema no convexo que es resoluble por métodos duales de descomposición jerárquica.	279
V	FV.10	Caso de un problema no convexo que no es resoluble por métodos duales de descomposición jerárquica.	279

=====

1.- Salvo indicación en contrario, todos los vectores utilizados se consideran como vectores columnas, y las funciones utilizadas diferenciables.

2.- El gradiente de una función $f(x)$, con respecto al vector x de sus variables independientes, se representa por

$\nabla_x f(x)$ y se considera como un vector fila.

3.- La expresión $\nabla_x g(x)$, con $x \in E^n$ y $E^n \xrightarrow{g} E^m$, significa el Jacobiano del vector de funciones g

$$\nabla_x g(x) = \begin{bmatrix} \nabla_x g_1(x) \\ \vdots \\ \nabla_x g_m(x) \end{bmatrix}$$

4.- x' , significa el vector fila transpuesto del vector columna x .

5.- $\dot{x}(t)$, ($x(t) \in E^n$) significa el vector de las derivadas primeras del vector $x(t)$ con respecto a la variable independiente t .

6.- La notación $x(\cdot)$ se utiliza en ocasiones para referirse globalmente a toda la trayectoria temporal, discreta o continua, del vector $x(t)$ entre dos instantes, inicial y final, prefijados.

=====

SINTESIS Y PERSPECTIVA

"L' économie, comme science
humaine, reste fondamenta-
lement politique..."

F. Mitterrand

"Models are to be used. Not
to be believed."

Theil

I. 1 PRESENTACION.

I. 2 SINTESIS.

I. 3 PERSPECTIVA.

Los más amplios y prometedores avances en la formulación y resolución práctica de modelos económicos, parecen situarse en la asimilación y desarrollo de la moderna teoría matemática del control óptimo.

La interacción entre ambos campos se desarrolla, a nuestro entender, en dos fases distintas por su contenido y cronología:

La primera de ellas cubre la década de los 60, aunque tenga sus raíces en el pionero artículo de Ramsey (R2). Su resultado fundamental es la teoría del crecimiento óptimo, en la que las condiciones necesarias enunciadas por el Principio del Máximo son aplicadas al análisis teórico de modelos hipersimplificados de crecimiento económico.

La utilización de los principios de optimización variacional constituyó, en su tiempo, una precursora iniciativa. Al mismo tiempo, los ingenieros de control seguían trabajando con los clásicos diagramas de Bode.

Después de las aportaciones de Pontryagin y Bellman, las exigencias prácticas (y el dinero) de la carrera del espacio producen el ingente desarrollo de las técnicas numéricas de estimación y control de los dos últimos tercios de la década (filtros de Kalman, teoría de las ecuaciones de Riccati, métodos del gradiente, etc...).

Los especialistas de tales temas, demasiado ocupados para preocuparse por potenciales aplicaciones económicas, circunscriben sus publicaciones a una literatura excesi-

vamente especializada, para atraer la atención de los economistas. La comunicación entre disciplinas no se produce. Falta de nuevos instrumentos, la literatura del crecimiento óptimo produce un conjunto de resultados tan elegantes como irrelevantes para guiar la acción práctica.

La interacción se produce de nuevo con el cambio de década. Los años transcurridos parecen indicar que una primera parte de esta segunda fase va a consistir en una masiva aplicación de los algoritmos de control en problemas económicos. Existen ya botones de muestra tan importantes como la literatura soviética de la "teoría de la planificación óptima" (D14,D15) y la aplicación por el Federal Reserve Board americano del regulador lineal-cuadrático en el análisis de medidas de política económica (A10,N1). En ambos casos, los campos de aplicación son determinados por el funcionamiento de los respectivos sistemas económicos y por los objetivos y medios que los organismos de control tienen encomendados, pero las técnicas aplicadas por unos y otros son esencialmente idénticas.

Las peculiares características de los sistemas económicos exigen modificaciones de métodos ya adquiridos y sugieren nuevos desarrollos. La formulación en tiempo discreto, la existencia de objetivos y centros de decisión múltiples y los métodos de descomposición necesarios para analizar sistemas de elevada dimensionalidad y estructura específica, constituyen ejemplos privilegiados. Atendiendo a estas necesidades, la teoría de los juegos diferenciales y los sistemas jerárquicos, o multiniveles, de control pueden convertirse en los instrumentos de una nueva y segunda parte.

amplia perspectiva así esquematizada, es el objetivo de este capítulo.

I. 2 SINTESIS

Nuestros dos grandes cuerpos de conocimiento previos son la teoría matemática del control óptimo y sus aplicaciones económicas clásicas. Ambos son tan extensos y definitivamente adquiridos que no es posible ni necesario su presentación resumida.

Limitándonos a recordar sus principales características, reducimos las referencias a algunas de las obras y "reviews" que contienen la extensa bibliografía que está fuera de nuestro objetivo e interés citar directamente.

Las relaciones entre los tres procedimientos de optimización dinámica: cálculo de variaciones (G1,B4), principio del Máximo de Pontryagin (P12,W1) y programación dinámica (B2,k3) han sido puestas de manifiesto por diversos autores (H4,D6). Tal relación es especialmente directa entre los dos primeros, considerándose al Principio del Máximo como una extensión del cálculo clásico de variaciones que supera muchas de sus limitaciones. Incluyendo a ambos bajo el nombre de métodos variacionales, el sistema canónico de Pontryagin es referido indistintamente como ecuaciones de Euler-Lagrange y la maximización del Hamiltoniano se corresponde con las condiciones de Weierstrass (P12,C8,H2), extendiendo su validez con independencia del dominio de valores admisibles de las variables de control.

Tal similitud no debe hacernos olvidar, que la ecuación de Euler se deriva bajo el supuesto de variaciones débiles, mientras que el Principio del Máximo se corresponde con variaciones fuertes. Ello quita validez, al popular dicho americano de que la Teoría del Control Optimo no es sino el "cáda-ver del cálculo de variaciones envuelto en la bandera soviética."

La programación dinámica presenta un enfoque y conduce a técnicas de cálculo diferentes. En régimen continuo conduce a la solución de una ecuación en derivadas parciales, la Hamilton-Jacobi-Bellman, y en régimen discreto a un procedimiento recurrente que rebasa todas las capacidades en memoria de acceso rápido actualmente disponibles cuando existen más de 4-5 variables de estado.

Por otra parte, sus hipótesis de diferenciabilidad de la función objetivo son más estrictas que las requeridas en el Principio del Máximo (P12, capítulo I). Aceptando las mismas, la ecuación de Hamilton-Jacobi resulta equivalente a la maximización del Hamiltoniano y la interpretación económica de las variables adjuntas es inmediata (P12, c8, A7).

El Principio del Máximo no es, ni ha tenido nunca la pretensión de ser, un procedimiento práctico de resolución salvo en algunos casos simples. Como las condiciones de Kuhn y Tucker, se limita a enunciar un conjunto de condiciones necesarias, que son suficientes bajo ciertas hipótesis de convexidad y positividad (K1, M2). El principio de optimalidad puede ser utilizado directamente como un procedimiento de resolución en problemas discretos de muy baja dimensionalidad pero la "curse of dimensionality" de la que se lamenta su autor anula tal capacidad en proble-

Ambos han servido como base a procedimientos numéricos de resolución que cubren la mayor parte de los 100 libros y 1500 artículos producidos por el desarrollo de la Teoría del Control Optimo en los últimos 25 años.

Algunos se refieren a la resolución del sistema de ecuaciones diferenciales con las condiciones en los límites divididas en los dos extremos de la variable de integración^{\$}. Otros a la reducción de las exigencias de la dimensionalidad, como el "state increment dynamic programming" de Larson (L1), a procedimientos directos como los métodos del gradiente, o en general, a métodos que parten de soluciones nominales que satisfacen solamente algunas condiciones de optimalidad y pasan a verificar el conjunto a través de un procedimiento iterativo.

Una visión sintética de este desarrollo, al que nos referiremos de nuevo en el capítulo 3, puede encontrarse en (P11,H1). Baste aquí recordar dos características importantes:

- la solución por el principio de máximo facilita, junto con la solución, los valores de las variables adjuntas cuya importancia en un análisis económico es igual a la de la información primal. Los métodos de programación dinámica no generan ningún tipo de información dual.
- la existencia de restricciones en las variables de estado complica la solución por el principio del máximo al tiempo que reduce la carga computacional en programación dinámica.

^{\$} Desde ahora referidos como TPBVP (Two Points Boundary Value Problems)

dice I resume los principales elementos del Principio del -
Máximo continuo.

Pero, en economía, los datos disponibles para alimentar un modelo aparecen como series de valores y no como funciones continuas del tiempo. Los controles disponibles no pueden ser variados sino a intervalos discretos. Por ello, uno de los más interesantes desarrollos para las aplicaciones económicas, es el de los métodos de control óptimo en tiempo discreto.

Las formulaciones del Principio del Máximo discreto son numerosas en la literatura desde 1964. Las primeras de ellas (J2,H1) están basadas en consideraciones geométricas que exigen ciertas hipótesis de regularidad. Las derivaciones subsiguientes (H8,R8) están basadas en la aplicación del teorema de Kuhn y Tucker expresando un problema discreto de control óptimo como un programa matemático de gran dimensión. Derivaciones detalladas pueden encontrarse en Albouy (A7,tomo II), Pyndick (P15, cap. II) y Wang (W1). El anexo II lo presenta siguiendo la línea de Rosen y Albouy.

En el campo de las aplicaciones económicas, el reconocimiento del interés de la teoría del control óptimo como medio instrumental de planificación es general: Smirnov (55), Brody (B6,B7), Kendrick (K6,K7), Fedorenko (F2), Chakavarty (C5), Fox-Sengupta (C7), Brolley (B8), Livesey (L9, L10), etc...

Sin embargo, hasta hace pocos años, todas sus aplicaciones se habían concentrado en la teoría del crecimiento óptimo y en problemas de estabilización a corto plazo. En el primer caso se desarrolló la línea empezada por Radner (R1) y

Instrumentalmente, la teoría del crecimiento óptimo se limita a aplicar el enunciado del Principio del Máximo para obtener información cualitativa acerca de la forma de las trayectorias de consumo y acumulación, óptimas en el sentido de una cierta funcional de utilidad colectiva, de un modelo de crecimiento neoclásico descrito por una o dos ecuaciones diferenciales. Una forma alternativa del objetivo, alcanzar una trayectoria de Von Neumann en tiempo mínimo, ha sido utilizada por Kurtz (K14), Stoleru (S10), Fedosev (F3).

La extensa literatura de la teoría del crecimiento óptimo ha hecho clásicos los nombres de sus autores: Arrow, Cass, Chakravarty, Koopmans, Mirrlees, Moristima, Phelps, Samuelson, Solow, Shell, Uzawa, etc ... Lejos de nuestra intención y espacio disponible el hacer un recuento de tal teoría y de sus derivaciones (problemas de estabilidad (K13), existencia (V5,M1), economías duales (D7,S8,N2), política fiscal etc. . Nos limitaremos a citar algunas referencias básicas (B9-S2) y los artículos de Dobell (D10) y Mirakhor (M9), que contienen una exhaustiva visión de la literatura. Desde el punto de vista de ingenieros de control existen los conocidos y más generales artículos de Athans (A10), Dobell y Ho (D9) y Dobell (D8).

Los diagramas de fases han sido ampliamente utilizados para describir las trayectorias óptimas y analizar los casos de control singular. El procedimiento es inutilizable cuando existen más de una variable de estado. Los casos de control singular han sido estudiados en Bryson (B5, capítulo 8) y Johnson (J1).

tituye la teoría de los "turnpikes" y el correspondiente énfasis en estados estacionarios ($C4, T2$) Rockafeller ha demostrado recientemente ($R4, R3$) que la tendencia de la trayectoria óptima a transcurrir la mayor parte del tiempo en las proximidades del estado estacionario y no alejarse de él sino para cumplir con las condiciones de transversalidad al final del período, es una propiedad general de los sistemas Hamiltonianos convexos en sus variables de estado y cóncavos en las adjuntas.

En problemas con horizonte infinito, en los que se esquivaba la necesidad de definir las condiciones finales de la trayectoria, Rockafeller demuestra que, en las anteriores condiciones, la trayectoria óptima es la que converge al estado estacionario.

Aunque matemáticamente justificada, la teoría de los "turnpike" no parece presentar mucho interés práctico y menor relevancia empírica. Problemas con horizonte infinito implican un cierto abuso de las capacidades de previsión y no son muy susceptibles de despertar entusiasmo entre los responsables de la política económica. Por otra parte, las perturbaciones aleatorias y cambios en las condiciones económicas generales impiden que un estado estacionario, sea óptimo teóricamente, sea susceptible de existir.

Es obvio que la capacidad descriptiva real de un sistema con los niveles de agregación utilizados en la teoría del crecimiento óptimo, es nula. Los intentos de formalización práctica ($M11, T8$) han sido escasos.

En el análisis de medidas de estabilización económica,

aplicaciones. Las obras de Phillips (P7,P8) y Theil (T6) están estrechamente relacionadas con los sistemas de regulación y especialmente con el regulador lineal-cuadrático (Apéndice IV). Este está en la base de las aplicaciones del tipo de - Pyndick (P15) y Chow (c6,c7) que forman parte, por cronología y contenido, de lo que hemos llamado segunda fase, y a los - que nos referimos en el capítulo IV.

Las aplicaciones en planificación económica, entendiendo por tal los procedimientos de optimización de la inversión en modelos dinámicos multisectoriales, son escasas pero importantes. Los únicos que no son conocidas son las de Kendrick (K6,K7), Martens (M4) y Dubobsiskii (D13). (Ver el "survey paper" de Alan Manne (M3)).

Son escasos porque precisan el empleo de algoritmos - numéricos para encontrar una solución concreta. Son importantes porque marcan la frontera (1970-71) entre las dos fases descritas y contienen todas las características de futuros desarrollos. A ellos nos referimos en el capítulo IV.

I. 3 PERSPECTIVA

Parte de los elementos fundamentales que constituyen un sistema matemático de planificación son de naturaleza política. Los objetivos que una sociedad se propone, o se le impone, alcanzar, las restricciones que definen las trayectorias aceptables y las condiciones finales en las mismas son otros tantos datos exógenos, previos y matemáticamente arbitrarios con respecto al posterior proceso de optimización y análisis. El concepto mismo de óptimo se define en el sentido que previamente determinan tales elementos.

te de los avances en los métodos matemáticos. En el estado actual de los mismos, los resultados que pueden obtenerse están lejos de ser exactos y directamente aplicables. Citando textualmente al profesor Eckaus en sus comentarios al artículo de A. Manne (M3)".... the image been created of the head of a planning commission pushing a button on a computer and having the next five-year plan roll out of the electric typewriter". Aunque esta información sea imposible de obtener de otra forma, el reverso de la imagen es la imposibilidad para los responsables de un plan de actuar con la sola base de una estrategia calculada numericamente.

Si ello fuera así, la información de tipo dual carecería de valor. Una vez que la trayectoria óptima esté única y perfectamente determinada resultan superfluos los indicadores que la balizan.

Puesto que este no es el caso, la información dual, las variaciones en el valor, pasa a dominar la escena. La valoración cuantitativa de las modificaciones en las condiciones exteriores y de las distintas alternativas de acción, constituye un elemento tanto o más importante que la búsqueda de una estrategia que es solamente óptima en condiciones muy específicas.

Aunque cualquier intento de analizar la adecuación entre acción y objetivos (optimización al fin y al cabo), solo puede realizarse mediante procedimientos analíticos a la altura de la dimensionalidad, dinámica y complejidad del sistema económico, un responsable político puede definir, por real ejercicio de su arbitraria voluntad, una trayectoria primal más o menos coherente. Pero le será mucho mas -

toria dual arbitraria.

Pontryagin, como Lagrange con sus multiplicadores, poco debia sospechar que las variables adjuntas iban a ser caracterizadas como un sistema de precios. Dada la importancia que poseen los resultados duales de un proceso de optimización dedicamos integralmente el capítulo II a la presentación unificada de las distintas interpretaciones económicas de variables adjuntas y multiplicadores.

Las limitaciones de la Teoría del crecimiento óptimo sugieren las nuevas líneas de aplicación de las técnicas de control óptimo. Clasificamos tales limitaciones en operativas y estructurales.

Limitaciones operativas

Se refieren a la necesidad de desagregación y a la introducción de elementos aleatorios. La segunda de estas partes se identifica con el desarrollo de la Teoría del control óptimo estocástico (B5,W3), de la que no nos ocuparemos en esta tesis, excepto por las consideraciones del punto IV. 2.

La desagregación pasa por el desarrollo de algoritmos capaces de resolver numericamente problemas de control con un número elevado de variables y restricciones en las variables de estado. El capítulo III lo dedicamos al estudio del desarrollo de algoritmos numéricos de resolución, cuyo proceso de cálculo simula la existencia de una única y todopoderosa autoridad en el sistema, constituyendo por tanto métodos globales de solución.

continuos y modelos de tiempo discreto en la formulación de modelos económicos dinámicos. Todos los algoritmos presentados se refieren a formulaciones en tiempo discreto.

La especial atención dedicada a los sistemas en tiempo discreto lo justificaremos en base a tres tipos de razones. De ellas, las dos primeras son en realidad causa y consecuencia impuestas por el uso precioso e insustituible del ordenador digital:

- 1) la discretización está en base de todos los procedimientos de cálculo con ordenadores digitales, de los que lo único que se puede obtener es una secuencia de valores con los que aproximar una función continua.
- 2) la solución de problemas continuos de control óptimo debe pasar por una u otra forma de discretización, aunque no sea más que para resolver aproximadamente los TPBVP, o integrar las ecuaciones de estado.
- 3) los problemas discretos o discretizados permiten la aplicación directa de las técnicas de programación matemática en espacios de dimensión finita.

Analizamos la alternativa entre métodos directos e indirectos de solución y entre programación matemática y control óptimo discreto. Las dos ópticas, la primal y la dual, con la que pueden ser abordados este tipo de problemas nos inducirá a comparar las ventajas e inconvenientes de sus respectivos procesos de cálculo y las hipótesis sobre las que se basan. Tal análisis comparativo es ilustrado en detalle mediante el desarrollo de una aplicación de algoritmos pri-

ma de control óptimo discreto (III.3), y comparándolo con un algoritmo que interpreta el método de los multiplicadores de Hestenes como un método dual a dos niveles (III.4.2). El papel jugado por las hipótesis de convexidad nos conduce a la consideración de los procedimientos de descomposición en el tiempo asociados al empleo de algoritmos duales puros tipo Lasdon-Tamura(III.4.3). Consecuentemente con el planteamiento adoptado, nos preocupará fundamentalmente el coste de cálculo y la precisión con el que los algoritmos diseñados, calculen, en el óptimo, los valores de las variables duales.

Una vez hayamos analizado las potencialidades prácticas de tales procedimientos de cálculo, pasamos a diseñar su aplicación a un modelo dinámico de planificación multisectorial a largo plazo, que exponemos en el capítulo IV.

Este modelo presenta una formulación no lineal.

Creemos

que la teoría del control óptimo está llamada a acompañar a la programación lineal en su monopolista papel como técnica de planificación económica. La programación lineal - tiene a su favor la universalidad de sus algoritmos y la generalización de los sistemas input-output como métodos de elaboración de datos macro-económicos. Todo lo que su utilización requiere es el uso de un generador de matrices.

En contrapartida:1) es de carácter intrínsecamente estático; 2) exige la adopción de hipótesis difícilmente - realistas; 3) el dominio de soluciones es solo cuasi-convexo implicando una deficiente estabilidad de los óptimos.El

difficulty which we economist will have in learning to use control theory is to remember what we have selectively forgotten about economics in our haste to make use of linear programming. We have forgotten that most economic phenomena are nonlinear rather than linear. Remember diminishing marginal utility and diminishing marginal product?".

La utilización generalizada de la teoría del control óptimo dependerá del grado de estandarización que alcancen sus algoritmos. En esta línea, los problemas lineales-cuadráticos son los mejores situados y de hecho la mayoría de las aplicaciones han sido de este tipo. En el apartado IV.2 nos referimos a las ventajas e inconvenientes relativos del modelo lineal-cuadrático, exponiendo las razones por las que no lo utilizamos en esta tesis.

Tampoco consideramos objetivos de la forma mínimo - tiempo, aunque sí seamos conscientes de las exigencias reales que traducen los problemas formulados con este objetivo. Como indica Brody (B6), el primer plan de la historia, el GOSPLAN para la electrificación de la URSS, está concebido como un problema de mínimo tiempo. Esta es la forma que adopta matemáticamente la urgencia en pasar de un estado inicial dado a otro considerado como preferible, y nadie dudaría en calificar de urgentes los problemas de muchas economías en desarrollo.

Limitaciones estructurales

La superación de las deficiencias estructurales no es una mera prolongación de la Teoría del crecimiento óptimo. Esta supone una funcional objetivo escalar y un único centro

prescinde de toda consideración acerca de la estructura - del sistema, de las relaciones de poder y conflictos entre sus partes.

En una economía capitalista ello equivale a negar la esencia misma de su funcionamiento. Pero incluso en una economía socialista planificada es preciso distinguir grados - de autonomía entre sus partes, objetivos y restricciones locales y niveles jerárquicos de coordinación y decisión. Citando a Fedorenko (F2)" the combination of centralized management and economic independence of the parts constitutes the essential road towards the improvement of the national economy".

La teoría del control óptimo generalizado TCOG (H6) considera funcionales objetivos vectoriales y varios centros de decisión. Por lo tanto engloba como casos particulares - los problemas de control óptimo y los juegos diferenciales (H6,s7). Sus aplicaciones en economía son por el momento escasas (L6,K12)pero la perspectiva es prometedora.

En la introducción al capítulo V (V.1), nos referimos a los trabajos (Pau (P3)) que presentan el proceso de planificación como la búsqueda de trayectorias de Nash en un juego diferencial entre los sectores de la economía. Tal interpretación, sería especialmente adecuada a las características del funcionamiento de una economía capitalista.

Por otra parte, la dimensionalidad de un problema no puede ser aumentada indefinidamente y seguir pretendiendo su solución global. El mismo fenómeno se presenta en programación lineal, en la que, a partir de un cierto número de va-

ción tipo Dantzig y Wolf (D3), Kornai y Liptak (K11)
Refiriéndonos en IV.4 a los límites de los métodos de solución global, dedicamos el capítulo V a los métodos jerárquicos de control multinivel de sistemas dotados de estructuras específicas (V.1, V.2) y al planteamiento descentralizado del modelo de planificación del capítulo IV (V.3)
Analizamos así, como el desarrollo de los métodos de descomposición jerárquica de los sistemas de control no solamente es provechoso numericamente, sino que se corresponde con las características estructurales de los sistemas económicos en una economía planificada.

Debido sin duda a su novedad, no conocemos ninguna aplicación concreta, como la elaborada, de la aplicación de los métodos de descomposición jerárquica de problemas de control óptimo, a modelos multisectoriales de planificación dinámica. En V.4 analizaremos las condiciones que debe reunir el problema para que el método sea aplicable, así como la importancia que presentan los subproductos de la información generada a lo largo de su proceso de cálculo, basada en las particulares características de los modelos económicos.

No es difícil suponer que la capacidad de tales métodos de descomponer un sistema complejo multidimensional en un conjunto de subsistemas interrelacionados de menor talla y coordinar las soluciones de los subsistemas para obtener el óptimo funcionamiento del conjunto, marca la línea de desarrollo de futuras aplicaciones económicas.

Finalmente, el capítulo VI resumirá las conclusiones a la que nos permite llegar el análisis de las cuestiones de las que el presente apartado constituye el hilo con-

La nomenclatura de los problemas que pasamos a analizar en el siguiente capítulo, está expuesta en los Apéndices I y II, por la primera página de los cuales se debe pasar necesariamente para proseguir la lectura. Tales apéndices constituyen una oportunidad, para el lector que lo desee, de recordar los enunciados del principio del Máximo continuo y discreto respectivamente.

A la lectura de los restantes apéndices se remite en el momento en el que su contenido está relacionado con el tema expuesto (III.1 y IV.2)

INTERPRETACION ECONOMICA DE LA TEORIA DEL CONTROL OPTIMO.

- "Prices are, first of all, a category connected with the optimal plan, where they must follow from, and not be established independently..."

Katsenelinboigen. (K4)

- II. 1 INTRODUCCION
- II. 2 OPTIMIZACION Y EQUILIBRIO.
- II. 3 SISTEMAS DE VALOR Y PERSPECTIVAS DE ACCION.
- II. 4 EL SISTEMA OPTIMO DE VALORACION.
- II. 5 LA DESCENTRALIZACION TEMPORAL DE LAS DECISIONES.
- II. 6 VALORACION DE LAS RESTRICCIONES.
- II. 7 VARIABLES ADJUNTAS Y ANALISIS DE SENSIBILIDAD.
- II. 8 CENTROS DE DECISION Y OBJETIVOS MULTIPLES.
- II. 9 ALCANCES PRACTICOS.

El propósito de este capítulo es presentar de forma unificada la interpretación económica del principio del máximo elaborada por distintos autores.

Los apartados 3 al 7 analizan el caso en el que existe un único centro de decisión y se encuadran en el marco general del análisis de sensibilidad frente a los distintos tipos de perturbaciones que pueden afectar a un sistema dinámico.

Un caso particular de las mismas, las que se traducen en variaciones de las variables de estado en un instante dado, presentan un doble interés :

- se corresponden con frecuentes acontecimientos reales (variaciones en los bienes de capital y fuerza de trabajo disponibles debidos a desastres naturales, averías, huelgas, etc...), o con las alternativas de elección de los responsables políticos (por ejemplo en los valores finales de los stocks de capital)
- su análisis permite interpretar las variables adjuntas como variables duales de las variaciones en las variables de estado y como un sistema óptimo de valoración (de precios) a través del cual se puede instrumentar la descentralización temporal de las decisiones.

Por la especial importancia de esta interpretación, se le dedican los puntos 3 al 6. Las interpretaciones relacionadas con otros tipos de perturbaciones se analizan en el apartado 7 como casos particulares de los teoremas enuncia-

dos por Peterson (P5) y Courtin-Routenberg (C10) en el análisis general de sensibilidad de problemas de control óptimo.

La conocida interpretación de $\Psi(0)$, primeramente expuesta por Dorfman (D12), como la variación marginal de la función objetivo con respecto a variaciones marginales en las condiciones iniciales del vector de estado $X(0)$ se obtiene como uno de estos casos particulares.

Su extensión a los valores correspondientes $\Psi(t)$, $X(t)$ de un punto arbitrario de la trayectoria óptima se obtendría inmediatamente, como lo hace Dorfman, mediante el uso implícito del principio de optimalidad.

Este recurso es cuidadosa y explícitamente evitado por Pontryagin en la presentación de sus teoremas (P12), y Albouy-Breton llegan al mismo resultado utilizando en su lugar los teoremas enunciados por Pallu de la Barrière (15,P). Los puntos 3 al 6 siguen fundamentalmente la exposición de dichos autores.

El punto 6 presenta el conflicto inherente a todo proceso de optimización dinámica como el equilibrio de un juego entre los espacios primal y dual. Estas interpretaciones, y las que se derivan en el apartado 7 del análisis de perturbaciones no marginales, no son sino la extensión al régimen dinámico de los conocidos conceptos resumidos en el punto 2.

El apartado 8 extiende la interpretación al caso de varios centros de decisión, enlazando con la proyección económica de la teoría de los juegos diferenciales.

· A su importancia teórica, las interpretaciones expuestas añaden una capacidad operativa que puede ser utilizada, tanto en la concepción de sistemas de planificación descentralizada (Armand (A9)) y génesis del correspondiente sistema instrumental de precios, como en el análisis post-optimal de los resultados y de las consecuencias de comportamientos sub-optimales.

Después de analizar estas potencialidades y sus límites, concluimos considerando que la Teoría del Control Optimo puede facilitar valiosos instrumentos de interacción y diálogo que ilustren a los responsables políticos de las consecuencias de sus acciones y facilite el análisis de alternativas.

En todo el capítulo utilizamos la notación definida en los Apendices I y II. El vector de estado representará las cantidades de bienes disponibles en un sistema económico en cada instante; las variables de control representarán la utilización que se haga de los mismos.

Es conocido que la propiedad de punto de silla del Lagrangiano de un programa matemático permite identificar el proceso de optimización con el equilibrio de un juego de suma nula entre los espacios primal y dual.

Dado el programa:

$$\begin{aligned} \text{Max } f(x) \\ g(x) \geq 0 \end{aligned}$$

sean x^* y μ^* los vectores de variables primales y duales (multiplicadores de Kuhn y Tucker) en el óptimo.

Por definición del Lagrangiano :

$$\mathcal{L}(x^*, \mu^*) = f(x^*) + \mu^* g(x^*)$$

$$\mathcal{L}(x^*, \mu) = f(x^*) + \mu g(x^*)$$

$$\mathcal{L}(x^*, \mu^*) - \mathcal{L}(x^*, \mu) = -\mu g(x^*) \leq 0$$

$$\boxed{\mathcal{L}(x^*, \mu^*) \leq \mathcal{L}(x^*, \mu)}$$

(1)

La búsqueda del vector μ^* :

$$\begin{aligned} \mu^* &= \text{Arg Min } \mathcal{L}(x^*, \mu) \\ \mu &\geq 0 \end{aligned}$$

constituye el problema dual del programa dado.

Supuesto que $f(x)$ y $g(x)$ son concavas y diferenciables,
es decir:

$$\begin{aligned} f(x) - f(x^*) &\leq \nabla f(x^*)(x - x^*) \\ g(x) - g(x^*) &\leq \nabla g(x^*)(x - x^*) \end{aligned}$$

la aplicación de la primera condición de Kuhn y Tucker, permite escribir :

$$\mu^* [g(x) - g(x^*)] \leq -\nabla f(x^*)(x - x^*) \leq f(x^*) - f(x)$$

es decir:

$$\boxed{\mathcal{L}(x, \mu^*) \leq \mathcal{L}(x^*, \mu^*)} \quad (2)$$

Las relaciones 1) y 2) conjuntamente se escriben :

$$\boxed{\mathcal{L}(x, \mu^*) \leq \mathcal{L}(x^*, \mu^*) \leq \mathcal{L}(x^*, \mu) ; \mu \geq 0} \quad (3)$$

o lo que es lo mismo :

$$\mathcal{L}(x^*, \mu^*) = \min_{\mu \geq 0} \left[\max_x (\mathcal{L}(x, \mu)) \right]$$

Supongamos que $\mathcal{L}(x, \mu)$ representa la función de ganancia del jugador que dispone de las variables x en un juego de suma nula que le opone al que dispone de las variables μ .

La condición para que este juego tenga solución no es otra que la expresada por la relación (3).

El jugador situado en el espacio primal se comportará pues de forma que :

$$\max_x \left[\min_{\mu} [\mathcal{L}(x, \mu)] \right] = \max_x [\mathcal{L}(x, \mu^*)]$$

Es obvio que no está haciendo más que resolver el programa matemático dado inicialmente.

Por su parte, el jugador situado en el espacio dual,

tratará de :

$$\min_{\mu \geq 0} \left[\max_x [\mathcal{L}(x, \mu)] \right] = \min_{\mu \geq 0} [\mathcal{L}(x^*, \mu)]$$

y al así hacerlo, resuelve el problema dual del programa dado, de acuerdo con la definición anterior.

Son también conocidas las interpretaciones económicas de los multiplicadores como:

- a) el incremento marginal del valor óptimo de la función objetivo resultante de variaciones marginales en el nivel de la restricción correspondiente.

Es decir, sea :

$$g_j(x) \geq b_j$$

$$\mu_j^* = \frac{\partial f(x^*)}{\partial b_j}$$

de lo que se deriva la conocida interpretación de los multiplicadores como un sistema de precios.

- b) para variaciones no marginales del nivel de las restricciones, los multiplicadores representan el límite superior de la variación media del valor óptimo de la función objetivo;

(P5)

$$\mu_j^* \geq \frac{f(x_1^*) - f(x_2^*)}{b_1 - b_2}$$

En la interpretación económica del Principio del Máximo se encuentra la extensión en economía dinámica de los conceptos así resumidos.

En el caso del Principio del Máximo discreto, la extensión

es inmediata puesto que (Apéndice II) las variables adjuntas son los multiplicadores de Lagrange asociados a las restricciones representadas por las ecuaciones de estado.

Al no poder efectuar justificadamente un paso al límite, el caso continuo exige una formulación mas laboriosa, pero en todos los casos, el contexto dinámico del problema enriquece conceptualmente los anteriores razonamientos, en la forma que analizamos a continuación.

II. 3 SISTEMAS DE VALOR Y PERSPECTIVAS DE ACCION.

El problema de la optimización en el tiempo de una función objetivo aditiva es fundamentalmente el de efectuar un arbitraje entre las consecuencias presentes y futuras de la cadena de decisiones instantáneas. Dado el estado del sistema $X(t)$, en un cierto instante t , toda decisión $u(t)$ origina un valor de $f_0(x, u, t)$ que es por si mismo una medida de las consecuencias inmediatas de $u(t)$. Al mismo tiempo, $u(t)$ afecta al estado del sistema y por lo tanto a las potencialidades que encierra.

Pero si la función objetivo asigna un valor al uso futuro de las potencialidades representadas por los valores de las variables de estado, el uso en si mismo solo queda definido cuanto una política de gestión o perspectiva de acción $u(\tau), \forall \tau \in [t, T]$ haya sido escogida.

A cada uno de los posibles conjuntos $u(\tau)$ corresponde un sistema de valor asociado a $X(t)$ con independencia de que $u(\tau), \forall \tau \in [t, T]$ sea o no la política óptima.

El valor que asociamos a $\chi(t)$ aparece como un valor de uso, definido como la variación que experimentaría la función objetivo entre t y T , debida a variaciones en la cantidad disponible de $\chi(t)$, si se siguiera una cierta política $\mu(\tau)$, $\forall \tau \in [t, T]$ de gestión del sistema.

En otras palabras, el valor de un bien (variable de estado) se define en función del uso (ley de control o política de acción) que pensemos hacer de él. Llamando :

$$\bar{J}[\chi(t), t] = \int_t^T f_0(\chi(\tau), \bar{\mu}(\tau), \tau) d\tau \quad (4)$$

una vez definida una específica política de gestión $\bar{\mu}(\tau)$, $\forall \tau \in [t, T]$, \bar{J} solo depende del punto $\chi(t), t$ a partir del cual se aplique dicha política.

De acuerdo con el concepto anterior de valor de uso, definimos un vector $\bar{p}(t)$ cada uno de cuyos componentes $\bar{p}_i(t)$ representa la variación de la función objetivo ante variaciones marginales de la cantidad disponible de dicho bien en el instante t , cuando se ha decidido seguir, a partir de t , una cierta política de acción $\bar{\mu}(\tau)$, $\forall \tau \in [t, T]$

$$\boxed{\bar{p}'(t) = \nabla_{\chi(t)} \bar{J}[\chi(t), t]} \quad (5)$$

Puesto que las variaciones de $\chi(t)$ son debidas a $\mu(t)$, a través de $p(t)$ podemos medir el valor de las transformaciones que $\mu(t)$ provoca y añadirlas a $f_0(x, u, t)$ para analizar su efecto total.

Derivando (4) :

$$f_0(x(t), \bar{\mu}(t), t) + \nabla_{x(t)} \bar{J}[x(t), t] \dot{x}(t) + \frac{\partial \bar{J}[x(t), t]}{\partial t} = 0$$

Usando (5), la ecuación anterior se escribe :

$$\boxed{f_0(x(t), \bar{\mu}(t), t) + \bar{p}'(t) \dot{x}(t) + \frac{\partial \bar{J}[x(t), t]}{\partial t} = 0} \quad (6)$$

Definiendo :

$$\begin{aligned} \bar{G} &= \bar{G}(x(t), \bar{\mu}(t), \bar{p}(t), t) = \\ &= f_0(x(t), \bar{\mu}(t), t) + \bar{p}'(t) \dot{x}(t) \end{aligned}$$

Obtenemos :

$$\begin{aligned} \boxed{\dot{x}'(t) = \nabla_{\bar{p}(t)} \bar{G}} \\ \nabla_{x(t)} \bar{G} = - \nabla_{x(t)} \frac{\partial \bar{J}}{\partial t} = - \frac{\partial}{\partial t} \nabla_{x(t)} \bar{J} = - \dot{\bar{p}}'(t) \\ \boxed{\dot{\bar{p}}'(t) = - \nabla_{x(t)} \bar{G}} \end{aligned} \quad (7)$$

El sistema (7) de ecuaciones diferenciales determina la evolución en el tiempo del vector $\bar{p}(t)$ asociado a la política de acción $\mu(t)$ en el intervalo $[t, T]$.

Tal sistema precisa de n condiciones en los límites:

- a) Si el estado final es libre y no existe función objetivo final :

$$\bar{J}[x(t), t] \rightarrow 0 \quad t \rightarrow T$$

es decir

$$\bar{p}(\tau) = 0$$

b) En el caso de que exista tal objetivo final:

$$\bar{p}'(\tau) = \nabla_{x(\tau)} \bar{J}[x(\tau), \tau] = \nabla_{x(\tau)} \Phi[x(\tau), \tau]$$

c) Si existen restricciones sobre el estado final

$$M[x(\tau), \tau] \geq 0$$

Modificando la función de evaluación para tenerlas en cuenta:

$$\begin{aligned} \bar{J}_1[x(t), t] &= \bar{J}[x(t), t] + \lambda' M[x(\tau), \tau] = \\ &= \int_t^T f_0(x(\tau), \bar{u}(\tau), \tau) d\tau + \lambda' M[x(\tau), \tau] \end{aligned}$$

Con lo que :

$$\boxed{\bar{p}'(\tau) = \lambda' \nabla_{x(\tau)} M[x(\tau), \tau]}$$

Puesto que $\bar{p}(t)$ representa el vector de los valores de uso de $x(t)$ asociados a $\bar{u}(t)$, $\dot{\bar{p}}(t)$ representa la variación en el tiempo de dichos valores, es decir, la depreciación sufrida por $x(t)$.

Dos sucesivas ampliaciones de los vectores $\bar{p}(t)$ y $x(t)$ hacen explícita :

a) la expresión de la conservación del valor del sistema :

Definiendo :

$$\bar{p}_{n+1}(t) = \frac{\partial \bar{J}[x(t), t]}{\partial t}$$

$$x_{n+1}(t) = t, \quad \dot{x}_{n+1}(t) = 1, \quad x_{n+1}(0) = 0$$

La ecuación (6) puede ahora escribirse :

$$f_0(x, \bar{u}, t) + \bar{p}'(t) \dot{x}(t) = 0$$

expresando que el incremento de utilidad originado por $u(t)$ y medido por f_0 es compensado por la pérdida de valor del sistema medida por $\bar{p}'(t) \dot{x}(t)$ y constituida por :

- la pérdida de valor "natural" debida al paso del tiempo. En este sentido decimos que $\bar{p}_{n+1}(t)$ es el valor del tiempo. Obsérvese que tal valor depende de la política de acción adoptada
- la pérdida de valor intrínseca debida a las variaciones $\dot{x}_i, i=1, n$

b) Una forma bilineal entre $\bar{p}(t)$ y $\dot{x}(t)$:

Definiendo :

$$\bar{p}_0(t) = 1$$

$$\dot{x}_0(t) = f_0(x, \bar{u}, t) \quad x_0(0) = 0$$

(6) se escribe :

$$\bar{p}'(t) \dot{x}(t) = \sum_{i=0}^{n+1} \bar{p}_i(t) \dot{x}_i(t) = 0 \quad (8)$$

por lo cual llamaremos a $\bar{p}(t)$ las variables duales de $\dot{x}(t)$

Hasta aquí hemos establecido un sistema de ecuaciones diferenciales, con sus correspondientes condiciones en los límites, que determinan la evolución de las variables $\bar{p}(t)$ como consecuencia de haber aplicado al sistema una ley de control arbitraria $\bar{u}(z), \forall z \in [t, T]$ a partir del estado

$X(t)$, tales variables $p(t)$ representan la valoración asignada a las cantidades de bienes representadas por $X(t)$, bajo el supuesto que de ellos se va a realizar el uso indicado por $\bar{u}(t)$.

Analicemos las características de este sistema de valoración, si la política $\bar{u}(t)$, $\forall t \in [t, T]$ fuera la política óptima.

II. 4 EL SISTEMA OPTIMO DE VALORACION.

En el caso en que la política de acción $\bar{u}(t)$ sea la óptima $u^*(t)$, llamando $p^*(t)$ el correspondiente sistema de valoración asociado, las anteriores ecuaciones se expresan :

$$\begin{aligned} G^* &= H^* \\ p^{*'}(t) &= -\nabla_{X(t)} H^* \\ \dot{X}^{*'}(t) &= \nabla_{p(t)} H^* \end{aligned} \quad (\text{Hamiltoniano óptimo})$$

Es decir, $p^*(t)$ obedece al mismo sistema de ecuaciones diferenciales que las variables adjuntas de Pontryagin, $\psi(t)$.

Puesto que tal sistema es lineal, admite una solución única para cada condición inicial. Basta pues demostrar la igualdad de $p^*(t)$ y de $\psi(t)$ para un valor de $t \in [0, T]$ para probar su identidad.

La condición de transversalidad del Anexo I define para $\psi(T)$ el mismo valor que las consideraciones del apartado anterior, asignan a $p^*(T)$. Podemos pues concluir que las variables adjuntas de Pontryagin representan las variaciones en la función objetivo que se podrían obtener debidas a variaciones marginales de $X(t)$, $t \in [0, T]$ si a partir de t se apli-

En otras palabras, $\psi(t)$ representa el valor de uso de $x(t)$ en el supuesto de su óptima utilización. Podemos pues, considerarlo como el precio máximo al que el centro de control estaría dispuesto a adquirir una unidad suplementaria de los bienes representados por las variables x . $\psi(t)$ Aparece así como el instrumento necesario para valorar las consecuencias de las transformaciones que $\mu(t)$ provoca. Su carácter de variables duales implica la existencia de un problema dual cuya primal sería el de maximización del Hamiltoniano.

Concretamente, el principio de Pontryagin define la búsqueda del vector óptimo de decisión $\mu^*(t)$ en cada instante t , como aquel que maximiza el Hamiltoniano construido con las funciones de valoración óptima $p^*(t) \equiv \psi(t)$ supuestas conocidas:

$$H[x^*, \mu^*, p^*, t] = \max_{\mu(t) \in U} H[x^*, \mu, p^*]$$

$$H[x^*, \mu^*, \psi^*, t] \geq H[x^*, \mu, \psi^*, t]$$

Suponiéndonos situados en un punto de la trayectoria óptima $[x^*(t), t]$, la ecuación (6) se particulariza:

=====

(*) De una forma mas rigurosa, la identidad del sistema óptimo de valoración y las variables adjuntos se demuestra apoyandose en los teoremas de Pallu de la Barrière acerca de las propiedades de $\psi(t)$. La ley de control $\bar{\mu}(t)$ se define en lazo cerrado con respecto al estado $\bar{\mu} = \bar{\mu}(x(t), t)$, la igualdad entre G^* y H^* satisfaciendose bajo la hipótesis de, no existencia de restricciones mixtas control-estado. (Albony(A5))

$$f_0(x^*(t), \bar{u}(t), t) + \nabla_{x(t)} \bar{J}[x^*(t), t] f(x^*(t), \bar{u}(t), t) + \frac{\partial \bar{J}[x^*(t), t]}{\partial t} = 0$$

$$f_0(x^*(t), \bar{u}(t), t) + \bar{p}'(t) f(x^*, \bar{u}, t) + \frac{\partial \bar{J}[x^*(t), t]}{\partial t} = 0 \quad (9)$$

El principio de Pontryagin permite escribir :

$$f_0(x^*(t), u^*(t), t) + p^{*'}(t) f(x^*, u^*, t) \geq f_0(x^*(t), \bar{u}(t), t) + p^{*'}(t) f(x^*, \bar{u}, t)$$

Sumando $\frac{\partial J^*[x^*(t), t]}{\partial t}$ a ambas partes de la anterior igualdad y observando que, según (6), el primer miembro se anula, obtenemos :

$$f_0(x^*, \bar{u}(t), t) + p^{*'}(t) f(x^*, \bar{u}, t) + \frac{\partial J^*[x^*(t), t]}{\partial t} \leq 0$$

La expresión anterior, junto con (9), permite escribir:

$$p^{*'}(t) f(x^*, \bar{u}, t) + \frac{\partial J^*[x^*(t), t]}{\partial t} \leq \bar{p}'(t) f(x^*, \bar{u}, t) + \frac{\partial \bar{J}[x^*(t), t]}{\partial t}$$

En el caso particular de que $\bar{u}(t) \equiv u^*(t)$, ello implica $\bar{J} = J^*$, y la desigualdad anterior se escribe :

$$p^{*'}(t) f(x^*, u^*, t) \leq \bar{p}'(t) f(x^*, u^*, t)$$

o lo que es lo mismo :

$$\boxed{H[x^*, u^*, p^*, t] \leq H[x^*, u^*, \bar{p}, t]}$$

desigualdad que expresa el problema dual :

$$\min_p [H[x^*, u^*, p]]$$

• El problema dual es pues el de encontrar entre todos los sistemas de valoración posibles, aquel que minimiza el Hamiltoniano construido con la política óptima $\mu^*(z), \forall z \in [t, T]$

El enunciado conjunto de los problemas primal y dual expresa que el proceso de optimización es el de la búsqueda del punto de silla del Hamiltoniano :

$$H[x^*, \mu, p^*, t] \leq H[x^*, \mu^*, p^*, t] \leq H[x^*, \mu^*, \bar{p}, t]$$

Tal estructura permite aplicar la interpretación dada en régimen estático y concebir el problema como el juego a suma nula entre dos adversarios dotados del control de las variables $\mu(t)$ y $p(t)$ respectivamente. El punto de equilibrio de tal juego está representado por la expresión anterior.

En el actual contexto dinámico, el " desdoblamiento ficticio de personalidad" del centro de decisión en dos centros antagónicos, ilustra el conflicto interno al que la preferencia por el presente y la vigilancia del futuro le someten en el momento de determinar sus decisiones instantáneas.

Bajo esta esquematización matemática, se diferencian y personalizan los elementos contrapuestos de cuyo enfrentamiento nacen nuestras decisiones en el tiempo.

Al resultado inmediato de la aplicación de $\mu(t)$, representado por la función objetivo instantánea $f_0(x(t), u(t), t)$, vienen a sumársele las consideraciones referentes al "cambio del valor" intrínseco del sistema.

Mientras la evaluación de las consecuencias inmediatas se realiza inmediatamente a través de $f_0(x(t), u(t), t)$, no.

existe, en el enunciado del problema, un medio explícito de evaluar operativamente las consecuencias futuras. El interés de las variables $\Psi(t)$ es el permitir evaluar tales consecuencias futuras, en el mismo momento en que se adoptan las decisiones que las motivan.

De esta forma, el equilibrio entre presente y futuro, puede ser analizado cuantitativa y operativamente. Dotados de un sistema dinámico de valoración de los estados del sistema, representado por $\Psi(t)$, es posible dividir el proceso de optimización global sobre todo el periodo considerado, en una serie de procesos de optimización instantáneos.

Es decir, dado el sistema de precios $\Psi(t)$, las decisiones instantáneas $\mu(t)$ pueden determinarse independientemente, como resultado de procesos locales de optimización llevados a cabo por centros autónomos que ejercen su pasajera autoridad en cada uno de los instantos de tiempo t . Tal proceso de descentralización es objeto de estudio detallado en el siguiente apartado.

II. 5 LA DESCENTRALIZACION TEMPORAL DE LAS DECISIONES.

El interés de la presentación detallada del diálogo entre los centros primal y dual es el de interpretar las variables $\Psi(t)$ como instrumentos de descomposición del proceso de optimización dinámica en una serie de procesos estáticos (determinación de $\mu(t), \forall t \in [0, T]$) garantizando a la vez la coherencia entre los objetivos instantáneos y el global.

Ello es válido tanto en régimen discreto como en continuo como lo muestra la exposición (Albouy (A7)), que desarrollamos conjuntamente en tiempo continuo y discreto:

- La preocupación por los resultados inmediatos:

$$\begin{cases} f_0(x(t), u(t), t) \\ f_0(x(k), u(k), k) \end{cases}$$

esta representada por el jugador primal, que dispone de las variables $\begin{cases} u(t) \\ u(k) \end{cases}$. Es el gerente del sistema. Por ejemplo, la compañía explotadora de un yacimiento.

El jugador dual, que fija el sistema de valoración $\begin{cases} \psi(t) \\ \psi(k) \end{cases}$, se preocupa de la explotación del sistema en el futuro y por lo tanto de las variaciones que en su valor provocan las decisiones primales. Es el propietario del sistema o, siguiendo el ejemplo, el país en cuyo territorio se encuentra el yacimiento.

Sea $\begin{cases} x(t) \\ x(k) \end{cases}$ el estado del sistema en un cierto momento del proceso. De acuerdo con la interpretación de $\psi(t)$ como un sistema de precios, el centro dual lo valora en

Pasado el intervalo siguiente $\begin{cases} \psi'(t)x(t) \\ \psi'(k)x(k) \end{cases}$ $\begin{cases} t, t+dt \\ k, k+1 \end{cases}$ lo valorará en:

$$\begin{cases} \psi'(t+dt)x(t) \\ \psi'(k+1)x(k) \end{cases}$$

Por lo tanto, está dispuesto a ceder el sistema en su estado $\begin{cases} x(t) \\ x(k) \end{cases}$ al centro primal a cambio de un "alquiler":

$$\begin{cases} [\psi'(t) - \psi'(t+dt)] x(t) \xrightarrow{dt \rightarrow 0} -\dot{\psi}'(t) x(t) \\ [\psi'(k) - \psi'(k+1)] x(k) \end{cases}$$

Pero este acuerdo sólo es válido en la medida en que el gerente devuelva el sistema en las mismas condiciones en que lo recibió, es decir, si no modifica $x(t)$ durante su gestión. Como que, por definición, sus decisiones afectarán a $x(t)$ en la forma definida por las ecuaciones de estado,

ambos centros llegarán a un acuerdo acerca de la contra-
partida de tales modificaciones.

Si el sistema es mejorado, $\left(\begin{cases} \dot{x}(t) > 0 \\ x(k+1) - x(k) > 0 \end{cases} \right)$ el propietario versará al gestor el valor que atribuye a tales modificaciones:

$$\begin{cases} \psi'(t) \dot{x}(t) \\ \psi(k+1) (x(k+1) - x(k)) \end{cases}$$

En caso contrario, el signo negativo del valor de la modificación indica que es el gerente quien deberá indemnizar al propietario.

El resultado neto que puede obtener el centro primal como resultado de su gestión $\begin{cases} \mu(t) \\ \mu(k) \end{cases}$ es pues:

$$\begin{cases} f_0(x(t), \mu(t), t) + \frac{\partial [\psi'(t) x(t)]}{\partial t} = f_0(x(t), \mu(t), t) + \psi'(t) \dot{x}(t) + \dot{\psi}'(t) x(t) \\ f_0(x(k), \mu(k), k) + \psi'(k+1) [x(k+1) - x(k)] + (\psi(k+1) - \psi(k))' x(k) \end{cases}$$

Mientras que el objetivo del centro dual es maximizar:

$$\begin{cases} - \frac{\partial [\psi'(t) x(t)]}{\partial t} \\ (\psi(k) - \psi(k+1))' x(k) - \psi'(k+1) (x(k+1) - x(k)) \end{cases}$$

Sobre el último termino de su función objetivo (el importe del "alquiler" del sistema) el centro primal no posee ningún control. Se limitará por lo tanto a:

$$\text{Max}_{\mu(t)} \left[f_0(x(t), \mu(t), t) + \psi'(t) \dot{x}(t) \right]$$

$$\text{Max}_{u(k)} \left[f_0(x(k), u(k), k) + \Psi'(k+1)(x(k+1) - x(k)) \right]$$

Es decir, a maximizar en cada instante el correspondiente Hamiltoniano definido en los Apéndices I y II.

Si en vez de considerar un intervalo $\begin{cases} \text{elemental} \\ \text{unitario} \end{cases}$, extendemos el análisis a un período $\begin{cases} t, t_1 \\ k, k+p \end{cases}$, las funciones objetivo representan fundamentalmente las mismas transacciones entre centros :

$$\begin{array}{l} \text{PRIMAL} \left\{ \begin{array}{l} \int_t^{t_1} f_0(x(z), u(z), z) dz + \Psi'(t_1)x(t_1) - \Psi'(t)x(t) \\ \sum_{j=k}^{k+p} f_0(x(j), u(j), j) + \Psi'(k+p)x(k+p) - \Psi'(k)x(k) \end{array} \right. \\ \\ \text{DUAL} \left\{ \begin{array}{l} \Psi'(t)x(t) - \Psi'(t_1)x(t_1) \\ \Psi'(k)x(k) - \Psi'(k+p)x(k+p) \end{array} \right. \end{array}$$

En ambos casos, la suma de las funciones objetivo de ambos centros se reduce a la verdadera función objetivo del problema original.

El Hamiltoniano aparece como la función objetivo que debe adoptar el gerente del sistema para maximizar el resultado neto de su gestión. Tal objetivo ficticio está contruido con el sistema de valoración que genera el centro dual. Puesto que los intereses de ambos son contrapuestos, el proceso solo alcanzará un compromiso cuando los valores de $\Psi(t)$ y de $u(t)$ sean tales que :

$$\underset{\mu}{\text{Max}} \left[\underset{\psi}{\text{Min}} H(x^*, \mu, \psi, t) \right] = \underset{\psi}{\text{Min}} \left[\underset{\mu}{\text{Max}} H(x^*, \mu, \psi, t) \right]$$

es decir, en el punto de silla representado por el valor óptimo del Hamiltoniano.

Una serie de resultados matemáticos aparecen ahora con un claro significado económico.

- a) Si la fecha final del proceso no esta fijada, debe ser aquella en la que las aportaciones globales al objetivo, (es decir el Hamiltoniano) de las decisiones óptimas sean nulas. Es decir, $H^*(T)=0$.
- b) Si el proceso es estacionario, el Hamiltoniano óptimo será constante a lo largo del proceso, puesto que siendo invariantes la estructura de evaluación y de comportamiento del sistema, no existe razón para modificar el equilibrio instantáneo alcanzado.

Cuando no existe tal invarianza, el centro de control trata de adaptarse a los cambios previstos variando el valor instantáneo del Hamiltoniano óptimo. Obsérvese que, puesto que se trata de un sistema determinista, tales cambios son perfectamente previsibles.

II. 6 VALORACION DE LAS RESTRICCIONES.

Las posibles restricciones sobre el estado final implican un sistema suplementario de valoración representado por sus multiplicadores λ asociados. Tal sistema de valoración está directamente relacionado con el que $\psi(\tau)$ representa

puesto que, a través de la condición de transversalidad, $\Psi(\tau)$ se define en función de λ y del gradiente de las restricciones.

Es interesante observar que en el caso de un estado final fijo (n restricciones igualdad), la aplicación de tal condición de transversalidad implica que ambos sistemas coinciden.

Sus significados económicos son los mismos, puesto que, en este caso, variaciones en el nivel de las restricciones es lo mismo que variaciones en las variables de estado. Pero la forma de enunciarlo asociada con λ (variación que hubiera sido posible obtener en la función objetivo global si la restricción no hubiera sido tan exigente) es mas operativa que la que $\Psi(\tau)$ representa (id.id. si el proceso se terminará disponiendo de una unidad mas de $x(T)$).

No podemos proyectar hacia el futuro el valor marginal de $x(T)$, ya que en T el proceso ha terminado. En su lugar lo proyectamos hacia el pasado, valorando así el condicionamiento que a lo largo de toda la trayectoria ha significado el deber alcanzar un cierto valor final de las variables de estado.

La indiferencia que expresa el hecho de no fijar a priori el valor de $X_i(\tau)$ indica que nuestra valoración implícita de los bienes representados por la componente i del vector de estado es nula. La condición natural en los límites, $\Psi_i(\tau)=0$, del sistema canónico se limita a indicar que el precio correspondiente es nulo en T .

. Por el contrario, el sistema de preferencia puede deli-

mitar, en forma mas o menos precisa, los valores finales admisibles del vector de estado. Ello implica aceptar una indeterminación en su valoración, expresada a través, o identificada con, los multiplicadores λ que forman parte de las $2n+q+1$ constantes del sistema canónico. En tal caso, sus precios solo pueden obtenerse como el resultado final del proceso de optimización.

Tal proceso genera pues, en los puntos límites del periodo de tiempo al que se extiende, las cantidades o los precios de las variables de estado, en función del planteamiento que se haya dado al problema. En lenguaje estrictamente económico, podemos decir que no es posible precisar, simultaneamente y a priori, precios y cantidades en un mismo instante del tiempo.

La valoración de las restricciones definidas a lo largo de la trayectoria, expresada a través de sus multiplicadores asociados, tiene en cuenta el doble efecto, inmediato y futuro, que seguiría a una modificación de las mismas.

En régimen estático el multiplicador $\lambda_i(k)$ asociado a la restricción $g_i(x(k), u(k), k) \leq G_i$ representa simplemente :

$$\lambda_i(k) = \frac{\partial f_0^*(x(k), u(k), k)}{\partial G_i}$$

En régimen dinámico, hay que añadir a dicho efecto instantáneo el que se induce en la variación de valor del sistema. Es decir, (en régimen discreto) :

$$\lambda_i(k) = \frac{\partial f_0^*(x(k), u(k), k)}{\partial G_i} + \frac{\partial [\Psi(k+1)x(k+1) - \Psi(k)x(k)]}{\partial G_i}$$

(A7)

• Ello implica reconocer que toda modificación de las restricciones instantáneas no solamente afecta a los resultados del periodo correspondiente , sino que se proyecta en el futuro modificando la valoración del sistema.

II. 7 VARIABLES ADJUNTAS Y ANALISIS DE SENSIBILIDAD.

Dos teoremas enunciados por Peterson (P5) y Courtin-Rootenberg (c10) explicitan el papel jugado por las variables adjuntas, y los multiplicadores asociados a los distintos tipos de restricciones , en el análisis de sensibilidad de problemas de control óptimo frente a una amplia y más compleja gama de perturbaciones que las que hasta ahora han sido consideradas.

La demostración de los siguientes teoremas puede encontrarse en las referencias citadas.

Dado el problema general de control óptimo del tipo y nomenclatura indicadas en el Apéndice I :

$$\underset{\substack{u(t) \\ t=0, T}}{\text{Max}} J(u(t)) = \int_0^T f_0(x(t), u(t), t) dt$$

$$\dot{x}(t) = f(x(t), u(t), t) \quad \text{ec. de estado}$$

$$M(x(t), u(t), t) \leq 0 \quad \text{restricciones instantáneas}$$

$$x(0) = x_0$$

$$x(T) = x_T$$

condiciones en los límites

$$\int_0^T L(x(t), u(t), t) dt \leq 0$$

restricciones integrales

considerémoslo como el miembro de la familia de problemas indexados por el parámetro ε , correspondiente a $\varepsilon = 0$:

$$\max_{\substack{u(t) \\ t=0, T}} J(u(t), \varepsilon) = \int_0^T f_0(x(t), u(t), t) dt$$

$$\dot{x} = f(x, u, t) + \varepsilon h(x, u, t)$$

$$M(x, u, t) + \varepsilon \phi(x, u, t) \leq 0$$

$$x(0) = a + \varepsilon \xi$$

$$x(T) = x_T + \varepsilon \eta$$

$$\int_0^T [L(x(t), u(t), t) + \varepsilon \ell(x(t), u(t), t)] dt \leq 0$$

en el que h , ϕ y ℓ son vectores de funciones continuas y ξ y η son vectores de constantes.

Supongamos que para cada valor de ε en un intervalo que incluya el valor 0, existen funciones $x(t, \varepsilon)$, $u(t, \varepsilon)$, $\mu(t, \varepsilon)$, y un vector de constantes con respecto al tiempo $\lambda(\varepsilon)_{(*)}$, que satisfacen las condiciones necesarias de Pontryagin y la condición de transversalidad.

$\mu(t, \varepsilon)$ es el vector de multiplicadores asociado a las restricciones instantáneas, por lo que

$$\mu(M + \varepsilon \phi) = 0 \quad \forall t \in [0, T]$$

=====

(*) Recuerdese que los multiplicadores asociados con restricciones de tipo integral son constantes. Ello es económicamente intuitivo. Una demostración matemática puede encontrarse en Valentine (v1).

analogamente

$$\lambda \int_0^T (L(x, u, t) + \varepsilon \ell(x, u, t)) dt = 0$$

Supuesto que $\lambda(\varepsilon)$ es diferenciable en $\varepsilon=0$, que :

$$\left. \frac{\partial x(t, \varepsilon)}{\partial \varepsilon} \right|_{\varepsilon=0} \quad \left. \frac{\partial u(t, \varepsilon)}{\partial \varepsilon} \right|_{\varepsilon=0}$$

son continuas

por intervalos para todo t y que, para todo $t \in [0, T]$

$$\left. \frac{\partial}{\partial t} \left(\frac{\partial x(t, \varepsilon)}{\partial \varepsilon} \right) \right|_{\varepsilon=0} = \left. \frac{\partial \dot{x}(t, \varepsilon)}{\partial \varepsilon} \right|_{\varepsilon=0}$$

Entonces :

Teorema 1

Si $u(t, \varepsilon)$ es la solución del problema de control óptimo correspondiente a un valor ε del parámetro y $u^0(t)$ es la solución correspondiente a $\varepsilon=0$,

$$\begin{aligned} & \frac{\partial}{\partial \varepsilon} [J[u(t, \varepsilon), \varepsilon]]_{\varepsilon=0} = \\ & = -\lambda(0) \int_0^T \ell(x^0, u^0, t) dt + \int_0^T \psi(t, 0) h(x^0, u^0, t) dt \\ & - \int_0^T \mu(t, 0) \phi(x^0, u^0, t) dt - \psi(T, 0) \eta + \psi(0, 0) \xi \end{aligned}$$

Con los supuestos adicionales:

- Concavidad del Hamiltoniano $H(x, u, \psi, t, \varepsilon)$ con respecto a x, u

- $\mu(t, \varepsilon) \geq 0$

Teorema 2 :

$$J[u(t, \varepsilon), \varepsilon] \leq J[u(t, 0), 0] + \varepsilon (\psi(0, 0) \xi - \psi(T, 0) \eta) \\ + \varepsilon \int_0^T [\psi(t, 0) h(x^0, u^0, t) - \lambda(0) \ell(x^0, u^0, t) - \mu(t, 0) \phi(x^0, u^0, t)] dt$$

El interés de ambos teoremas es el de permitir interpretar las variables adjuntas en el contexto de la relación entre modificaciones estructurales del sistema y las variaciones engendradas como consecuencia en los valores óptimos de la función objetivo. El segundo de ellos permite analizar la respuesta a modificaciones no marginales.

Las perturbaciones que ahora consideramos son más generales y profundas. Están inmersas en las propias ecuaciones de evolución del sistema y alteran los niveles de las restricciones instantáneas e integrales a las que lo sometemos. Por estas razones es mayor su importancia en un análisis de comportamientos subóptimos, errores en la formulación del modelo o cambios en las exigencias de tipo político introducidas a través de las restricciones.

La información que variables adjuntas y multiplicadores contienen, utilizada en la forma que los teoremas anteriores establecen, permite instrumentar metódicamente el diálogo entre los resultados numéricos del modelo, sus imperfecciones, las variaciones en la realidad que trata de describir y nuestras cambiantes alternativas y preferencias.

Particularicemos los teoremas y consideraciones anteriores en el contexto de un modelo simple de crecimiento económico como el propuesto por Solow y Dorfman (*) (D12)

$$\text{Max}_{c(t)} \int_0^T e^{(n-e)t} u(c(t)) dt$$

$$\dot{K} = f(k) - c - (n+\delta)k$$

$$K(0) = k_0$$

$$K(T) = k_T$$

$$c(t) \geq a k(t)$$

Suponiendo que todas las perturbaciones son nulas exepcto las que afectan a las condiciones iniciales y que en estas

$$\xi=1 \quad K(0) = K_0 + \xi$$

El primer teorema establece :

$$\frac{\partial}{\partial \xi} [J[c(t, \xi), \xi]]_{\xi=0} = \psi(0, 0)$$

Es decir, si dispusieramos de una "unidad" mas de capital en el instante inicial, $\psi(0)$ mediría la variación que a lo largo del proceso experimentaría el valor de la función objetivo como consecuencia de su óptima utilización. Este resultado fue derivado por otro procedimiento por Dorfman (D12) y no es sino la particularización en un cierto instante (el inicial) del análisis que nos ha permitido identificar

- (*) Nomenclatura: $K(t)$:capital per cápita (única variable de estado)
 $c(t)$:consumo per cápita (" " "control")
 n :ritmo de crecimiento de la población
 e :tipo de actualización
 δ :tipo de depreciación del capital

las variables adjuntas con un sistema de precios.

Supongamos ahora que solamente son no nulas las perturbaciones inmersas en las ecuaciones de evolución, es decir, las variaciones en la ley de acumulación de capital .

Si después de haber realizado el cálculo y encontrado los valores correspondientes de $\Psi(t)$, resultará que nuevas estimaciones indicarán una ligera modificación en la forma de la función de producción que hiciera creer que $f^*(k)$ es mas apropiada que $f(k)$, la variación en los valores alcanzables de la función objetivo en el caso de una política óptima nos vendría dada por :

$$\int_0^T \Psi(t,0) (f^*(k) - f(k)) dt$$

Con frecuencia, las politicas "optimas" de desarrollo generadas por distintos modelos teóricos muestran una trayectoria politicamente inaceptable por los bajos niveles de consumo que determinan para los primeros periodos del proceso. Un ejemplo interesante es el ya clásico modelo de Stoleru (89) aplicado a la economía argelina.

En tal caso, se introducen nuevas restricciones en el modelo para generar una trayectoria (óptima también en su correspondiente contexto) que garantice un nivel mínimo de consumo, expresado, por ejemplo, como una proporción a de la capacidad productiva.

El valor de a, o el de otro procedimiento empleado para fijar cuantitativamente tal nivel de consumo, es una elección política. El modelo no puede hacer mas que aceptarla como uno de sus elementos definitorios.

Al así hacerlo genera necesariamente, en el consiguiente proceso de cálculo, una información que es fundamental para ilustrar la decisión precedente : ¿ cual sería la variación del valor óptimo de la función objetivo ante ligeros desplazamientos del nivel garantizado de consumo? En nuestro sencillo ejemplo, si las perturbaciones representan variaciones en la constante \underline{a} , la respuesta es

$$-\int_0^T \mu(t) k(t) dt$$

expresión en la que $\mu(t)$ es el multiplicador asociado a $C(t) - a k(t) \geq 0$ y $k(t)$ es la trayectoria óptima asociada al valor original de \underline{a} .

Uno de los objetivos fundamentales de un modelo operativo de planificación es el de apreciar las consecuencias de los valores escogidos en el stock final de capital $K(T)$ así como el de testar la sensibilidad, frente a sus variaciones, de los resultados obtenidos.

El teorema 1 de Peterson nos sigue informando que toda exigencia suplementaria en la situación final se traduciría en una disminución del óptimo alcanzable durante el periodo por un valor $\psi(T)$. Aquí son aplicables las consideraciones expuestas en el apartado II. 6 acerca de la valoración de las restricciones.

Todas las perturbaciones hasta ahora consideradas son de tipo teórico marginal. Todas las interpretaciones expuestas se reducen al análisis de variaciones marginales en una cantidad frente a modificaciones marginales en otras.

Cuando las modificaciones en los niveles de las restric-

ciones no son marginales, sino finitos, al apartado II.1 recuerda que los multiplicadores representan límites superiores en la variación relativa del valor óptimo del objetivo. El teorema 2 de Peterson nos permite extender tal interpretación a los problemas dinámicos.

Ello implicaría que si en vez de la situación inicial $K(0) = K_0$, dispusiéramos de una cantidad inicial de capital $K_0 + K_0^s$:

$$\psi(0) \geq \frac{J[K_0 + K_0^s] - J[K_0]}{K_0^s}$$

en donde $J[\cdot]$ representa el valor óptimo de la función objetivo alcanzable en los distintos casos de dotación inicial de capital.

Analogamente, si la proporción a de capital acumulado que se establece como nivel mínimo de consumo pasara a ser $a + a^s(t)$ la máxima variación relativa en el valor óptimo previamente calculado de la función objetivo sería:

$$-\int a^s(t) \mu(t) K(t) dt$$

Hasta aquí, un análisis teórico nos ha permitido interpretar los subproductos de la optimización, variables adjuntas y multiplicadores, como un sistema de valoración de las cantidades de bienes disponibles y de las modificaciones en la estructura y/o comportamiento del sistema analizado. A lo largo de toda la discusión, hemos supuesto que todas las decisiones dependían de un único centro dotado de un único objetivo. Esta hipótesis traduce difícilmente la realidad de los sistemas económicos. El apartado siguiente extiende las anteriores interpretaciones al caso en el que

tal hipótesis es abandonada.

II. 8 CENTROS DE DECISIÓN Y OBJETIVOS MÚLTIPLES.

Cualquiera que sea la organización del sistema económico, la existencia de centros de decisión múltiples es una de sus características esenciales. Dotados de mayor o menor autonomía y poder, guiados por objetivos mas o menos contrapuestos, el conflicto entre centros se suma al conflicto entre presente y futuro de cada uno de ellos.

El conflicto entre centros puede representarse a través de las ecuaciones de estado o de sus funciones objetivo. Las primeras expresan como el vector de estado se modifica bajo las acciones de control ejercidas por los distintos centros. La estructura de evaluación de cada centro expresa su sensibilidad frente a su posición relativa en el conjunto, así como la valoración que asigna a las acciones de los demás.

Representando por $\mu^j(t)$ el vector de variables de control del centro j , las ecuaciones de estado y las funciones objetivo en un juego diferencial entre N jugadores se escriben: (K12, I2)

$$\dot{X}(t) = f(X(t), \mu^1(t), \dots, \mu^j(t), \dots, \mu^N(t), t), X(0) = X_0$$

$$\begin{aligned} \underset{\mu^1(\tau)}{\text{Max}} J_1(X(t), \mu^1(t), \dots, \mu^j(t), \dots, \mu^N(t)) &= \text{Max} \left[\Phi_1(X(\tau), \tau) + \int_0^T f_0^1(X(t), \mu^1(t), \dots, \mu^N(t)) dt \right] \\ &\vdots \\ \underset{\mu^N(\tau)}{\text{Max}} J_N(X(t), \mu^1(t), \dots, \mu^j(t), \dots, \mu^N(t)) &= \text{Max} \left[\Phi_N(X(\tau), \tau) + \int_0^T f_0^N(X(t), \mu^1(t), \dots, \mu^N(t)) dt \right] \end{aligned}$$

En los apartados anteriores hemos visto como a cada política de acción $\mu(t)$, instrumentada por un único centro de decisión, le corresponderá un sistema de valoración asociado. Los "indicadores de valor" $\Psi(t)$ de las cantidades de bienes $X(t)$ dependen del uso futuro que de ellos se proponga hacer dicho centro. Desde el instante en que existen otros entes con capacidad autónoma de "hacer uso" de tales bienes (matemáticamente hablando, de influir a través de $\mu^j(t)$ sobre $\dot{X}(t)$, en la persecución de sus propios objetivos), el sistema de valoración $\Psi(t)$ deja de ser único. A cada variable de estado $X_i(t)$ le corresponden tantos sistemas de valoración $\Psi_i^j(t), (j=1, \dots, N)$, como centros de decisión existan. En lugar del vector $\Psi(t)$ existe ahora una matriz $\Psi(t)$ de dimensiones $(n \times N)$.

Cada uno de sus elementos $\Psi_i^j(t)$, representa la utilidad marginal, en el sentido de su propia función objetivo, que para el centro j posee el bien i en el instante t , bajo el supuesto que, tanto él como los demás centros, van a aplicar unas ciertas políticas de acción.

En un juego estático, ello implicaría que los sistemas de valor son función de las actitudes adoptadas por los jugadores. Dado que el proceso posee un desarrollo temporal, la determinación de $\Psi(t)$ depende no solamente de las actitudes adoptadas, sino de la generación de información disponible a cada centro.

En los tipos de sistemas deterministas transitivos que estamos estudiando, la información generada se resume en los valores de las variables de estado $X(t)$. Si los distintos centros de decisión tienen acceso a tal información, establecerán sus estrategias de acción, no solamente como fun-

ciones del tiempo, sino como funciones del estado que el sistema alcance bajo la acción conjunta de todos ellos. Es decir, $\mu^j \equiv \mu^j(t, x(t))$. Son los llamados sistemas de control "en lazo cerrado" (closed-loop) como contrapartida de los sistemas en lazo abierto $\mu^j = \mu^j(t)$. (Citron (c8)).

El tipo de juego menos relevante para los fenómenos económicos es el duelo. No solamente por su reducido número de jugadores sino porque la absoluta contraposición de intereses es raramente el caso en las relaciones económicas. En el caso del duelo, es evidente que las dos columnas de la matriz $\Psi(t), (n \times 2)$ son iguales y de signo contrario, expresando la apreciación simétrica que ambos jugadores poseen de las modificaciones en el sistema.

Los juegos a suma no nula entre N jugadores representan una mejor aproximación a los fenómenos económicos. Distintos tipos de "solución" pueden ser definidos dependiendo fundamentalmente de la actitud más o menos cooperativa de los jugadores y de la información a la que tienen acceso.

Esta puede ser tan pobre que induzca a los distintos centros a creer que un antagonismo total existe entre ellos. En tal caso, las trayectorias minimax se obtendrían como el resultado de N duelos diferenciales, oponiendo cada uno de ellos sucesivamente a un centro con la coalición de los $N-1$ restantes.

Las ecuaciones definitorias de los sistemas de valoración $\psi^j(t)$ varían según que el proceso se desarrolle en lazo abierto (Albouy (A7)) :

$$\dot{\psi}^j(t) = -\nabla_{x(t)} H_j$$

$$\psi^j(\tau) = \nabla_{x(\tau)} \Phi_j(x(\tau), \tau) \quad (j=1, \dots, N)$$

o lazo cerrado :

$$\dot{\psi}^j(t) = -\nabla_{x(t)} H_i - \sum_{i=1}^N (\nabla_{u^i(t)} H_j) (\nabla_{x(t)} u^i(t))'$$

$$\psi^j(\tau) = \nabla_{x(\tau)} \Phi_j(x(\tau), \tau) \quad (j=1, \dots, N)$$

Expresiones en las que $H_i (i=1, N)$ son los Hamiltonianos de los distintos centros de decisión:

$$H_i = f_0^j(x(t), u^1(t), \dots, u^j(t), \dots, u^N(t)) + \psi^j(t) f_j(x(t), u^1(t), \dots, u^N(t), t)$$

Ambas expresiones se refieren al caso en el que no existen restricciones. En tales condiciones y en el óptimo, se debe verificar que :

$$\nabla_{u^j(t)} H_i = 0 \quad (\forall j=1, \dots, N)$$

con lo que se establece la igualdad entre las trayectorias minimax en lazo abierto y cerrado. Las expresiones se complican mas con la existencia de restricciones pero el mismo resultado sigue siendo válido (Albouy (A7)).

En otros términos, el acceso a mayor información no altera los sistemas de valor generados por cada centro siguiendo una actitud decididamente antagonista. Sin embargo, lo que matemáticamente no se puede establecer, pero si es presumible que ocurra, es el abandono de tal actitud en un sistema controlado en lazo cerrado. Sabido es que si todos

los jugadores siguen sus respectivas estrategias de defensa, el resultado obtenido será distinto del esperado por cada uno de ellos. Una correlación inteligente entre tal fenómeno y la información contenida en el vector de estado convencerá a los centros de que su supuesto antagonismo no existe, y les inducirá a actitudes de compromiso y cooperación.

Supongamos la existencia de tal capacidad de raciocinio pero admitamos que no es lo suficientemente fuerte como para conducir a los distintos centros a una acción concertada. En tal caso un conjunto de estrategias de interés, suponiendo su existencia, son aquellas que conducen a las trayectorias de equilibrio de Nash. Estas tienen la misma definición que en la teoría estática de juegos :

$$\mu^*(t) = \{ \mu^{1*}(t), \dots, \mu^{N*}(t) \}$$

es una estrategia de equilibrio de Nash si :

$$\forall i = 1, \dots, N$$

$$J_i(\mu^*(t)) = \max_{\mu_i} J_i(\mu^{1*}(t), \dots, \mu^{(i-1)*}(t), \mu^i(t), \mu^{(i+1)*}(t), \dots, \mu^{N*}(t))$$

Ningún jugador puede alcanzar mejores resultados abandonando unilateralmente su estrategia de Nash. Por supuesto, una acción concertada podría conducir a estrategias eficientes tipo Pareto gracias a las cuales podrían ser obtenidos mejores resultados para todos los centros. En general, las trayectorias de Nash no pertenecen al conjunto de trayectorias eficientes.

Si los conjuntos de estrategias de mejor respuesta de

Los distintos centros no poseen una intersección vacía, no existen trayectorias de Nash. Condiciones bastante restrictivas para su existencia han sido estudiadas por Varaiya (v2).

Case (8,99) presenta condiciones necesarias generales para que una cierta estrategia $\mu^j(t), j=1, N$ sea una estrategia de equilibrio de Nash. Utiliza dos aproximaciones, una siguiendo procedimientos de programación dinámica y la otra empleando el principio del máximo. Estos últimos resultados han sido extendidos a estrategias definidas en lazo cerrado por Starr y Ho (H6).

Caso de existir , las trayectorias de Nash presentan una propiedad importante desde el punto de vista de la interpretación económica: las trayectorias en lazo abierto no coinciden con las trayectorias en lazo cerrado.

Tal resultado no es intuitivo en una primera aproximación. En un problema determinista de control óptimo las "leyes" $\mu(t)$ y "funciones" $\mu(x(t), t)$ de control pueden derivarse mutuamente una de otra. Puesto que, teóricamente, la búsqueda de las estrategias de Nash equivale a resolver N problemas paramétricos de Control Óptimo y buscar la intersección de sus soluciones, y puesto que tal proceso puede realizarse a priori ya que no existen "inputs" impredecibles en el sistema, parece que los controles resultantes debieran seguir gozando de la misma propiedad.

Sin embargo, la ruptura que representa el acceder de uno a varios centros de decisión invalida tal razonamiento. En la búsqueda de trayectorias de Nash el Hamiltoniano de

cada centro aparece como :

$$H_i = f_0^i(x(t), u^{1*}(t), \dots, u^i(t), \dots, u^{N*}(t)) + \psi^j(t) f^j(x(t), u^{1*}(t), \dots, u^i(t), \dots, u^{N*}(t))$$

En ausencia de restricciones, las condiciones necesarias de una trayectoria de Nash en lazo abierto establecen las ecuaciones definitorias de $\psi^i(t)$ como :

$$\dot{\psi}^i(t) = -\nabla_{x(t)} H_i$$

En lazo cerrado :

$$\dot{\psi}^i(t) = -\nabla_{x(t)} H_i - \sum_{\substack{j=1 \\ j \neq i}}^N \left(\nabla_{u^{j*}(t)} [H_i] \right) \left(\nabla_{x(t)} [u^{j*}(t)] \right)'$$

De la única forma en que ambos sistemas de valoración pueden coincidir es anulándose el sumatorio de la expresión anterior. Ello ocurrirá cuando :

- no haya retorno de información (lazo abierto)
- se trata de un duelo (en cuyo caso las trayectorias de Nash en lazo abierto y cerrado coinciden con la minimax)
- Cuando $N=1$ (problema de control óptimo).

En ausencia de estos tres casos particulares el sumatorio no será nulo en general, puesto que los términos

$$\nabla_{u^{j*}} H_i \neq 0 \quad (\forall j \neq i)$$

Ello es debido a que, a lo largo de una trayectoria de Nash, cada centro minimiza su Hamiltoniano con respecto a sus propias variables de control pero no puede evidentemente hacer lo mismo con respecto a aquellas que dependen de los demás centros.

Si existiera un conjunto de estrategias de control $u^j(t)$

• $i=1, \dots, N$ que maximizará simultáneamente todos los Hamiltonianos, el problema se reduciría a N problemas de control óptimo en los que cada centro tendría a su disposición los N vectores de control. En este caso, la "trayectoria de Nash" obtenida sería también "eficiente" puesto que todos los centros escogerían la misma.

Este caso degenerado aparte, la sensibilidad $(\nabla_{x^i(t)} H_i \neq 0)$ del objetivo que el centro i debe adoptar a efectos de analizar una optimización temporal, frente a las estrategias seguidas por los demás centros, encierra dos importantes consecuencias:

- a) Computacionalmente, las ecuaciones necesarias de la trayectoria de equilibrio se convierten en un sistema de ecuaciones diferenciales en derivadas parciales extremadamente difícil de resolver.
- b) Interpretativamente, los sistemas de valor $\psi^i(t)$ de cada centro dependen explícitamente de las reacciones de los demás centros ante variaciones en la trayectoria.

Analicemos tal mecanismo de dependencia : Las variables adjuntas se interpretan en el análisis de sensibilidad frente a perturbaciones. Concretamente, $\psi^i(t)$ representa la variación marginal de la función objetivo del centro i con respecto a una perturbación δx_i de una componente del vector de estado.

Pero cualquier variación del vector de estado, δx , inducirá cambios en las estrategias de los demás centros si estas están definidas en lazo cerrado. Tales variaciones son de

magnitud $(\nabla_{x_i}(t) \cdot \nabla_{x_i}(t)) / \|\nabla_{x_i}(t)\|$, y, puesto que $\nabla_{x_i} H_i \neq 0$, se traducen en variaciones suplementarias del objetivo del centro i .

Las ecuaciones que definen $\dot{\Psi}^j(t)$ indican simplemente que tal efecto de retorno debe ser tenido en cuenta en la valoración de sus actuaciones para cada centro. Suponiendo que una solución pueda ser encontrada, el conocimiento por cada centro de su columna correspondiente en la matriz $\Psi(t)$ le informa del valor que para él representa cualquier intento de modificación del vector de estado, incluyendo en este las reacciones de los demás centros.

Es evidente que la matriz $\Psi(t)$ tendrá un buen número de elementos nulos. En cada columna, habrá tantos ceros como variables de estado estén ausentes de la función objetivo del centro correspondiente. Ello indica que el centro en cuestión es indiferente frente a variaciones en tales variables de estado. Pero no implica necesariamente que este desprovisto de la posibilidad de actuación sobre ellas y por lo tanto de influir en el sistema de valoración de los demás centros. En otras palabras, si por una parte los campos de valoración de los distintos centros no son disjuntos, sus dominios de acción no son generalmente idénticos a respectivos campos de valoración. De aquí, todo el interés de llegar a soluciones que rebasen el compromiso para llegar a la cooperación.

La búsqueda de trayectorias de Pareto consiste fundamentalmente en la resolución de un problema de control óptimo con una función objetivo vectorial. A este respecto, son de fundamental importancia los resultados de Da Cunha y Polak (D1). Desde un punto de vista de interpretación eco

nómica, la matriz $\Psi(t)$ se reduce de nuevo a un vector, indicando el sistema de valor para la "entente" de los N jugadores de las variaciones en el vector de estado. Pero tal sistema de valoración no aporta, en sí mismo, ninguna información para escoger entre las distintas trayectorias eficientes posibles.

Acabamos aquí este breve resumen de la extensión al caso de varios centros de decisión de la interpretación de las variables adjuntas, para considerar, en el siguiente apartado, los condicionamientos prácticos y el alcance real de tales interpretaciones teóricas.

II. 9 ALCANCES PRACTICOS

Los puntos anteriores han tratado de analizar, desde una perspectiva económica, la información contenida en variables adjuntas y multiplicadores generados en el proceso de optimización de sistemas dinámicos. Su relevancia ha sido ilustrada mediante ejemplos voluntariamente simples. Siempre a nivel teórico, la misma relevancia informativa se extiende automáticamente a modelos complejos multidimensionales.

Desde un punto de vista práctico, las posibilidades e interés de su utilización, dependen de una serie de condicionamientos. Para analizarlos, adoptemos el conjunto de hipótesis más favorables. Sucesivas relajaciones de las mismas nos permitirán observar como varía disparmente el interés de algunas potenciales utilidades de la información dual.

Supongamos que nos atrevemos a aceptar explícitamente los supuestos que, consciente o inconscientemente, están en la base de muchos modelos económicos:

a) Las posibilidades de cálculo son ilimitadas.

Cualquier modelo matemático puede ser resuelto independientemente de su estructura y dimensionalidad.

b) La capacidad de modelización es perfecta.

Los modelos resultantes describen exacta y precisamente los mecanismos y preferencias del sistema descrito.

c) El sistema es invariante.

Las preferencias expresadas y los mecanismos descubiertos son definitivos y/o sus cambios perfectamente previsibles.

En tan ideales condiciones, la medida de "variaciones en el valor" frente a perturbaciones y alternativas carece de interés desde el momento en que éstas no se producen. Por el contrario, la capacidad de descentralización temporal de las variables adjuntas resulta primordial puesto que bajo el sistema de precios que representan, el comportamiento resultante de los agentes económicos se sitúa sobre la trayectoria primal óptima. Una planificación por las cantidades, centralizada, costosa y lenta, puede ser teóricamente substituida por una planificación por los precios, descentralizada, impersonal y automática.

Tal sistema de precios posee toda la significación de esta categoría económica: un sistema óptimo de valoración. Tales calificativos presuponen ciertamente un conjunto de elecciones previas: definición de aquello a lo que las partes integrantes del sistema (objetivo,s) asignan valor y el

cómo, así como el conjunto de situaciones que consideran como aceptables. Ambos elementos revelan decisiones de tipo político a través de las cuales se introduce en el funcionamiento del sistema económico un importante elemento de subjetivismo.

Pero tales sistemas de valor no se derivan automáticamente de la simple aceptación como datos de unas opciones políticas. Solamente un análisis de las distintas trayectorias posibles y su clasificación mediante la función objetivo, puede generarlos. En este punto crucial, se sitúa todo el interés de los métodos aplicados de optimización y control en economía.

Abandonemos el mundo ideal en el que nos habíamos situado. No hace falta decir que los modelos matemáticos distan mucho de reflejar la complejidad real de los sistemas económicos. Y estos están, afortunadamente, lejos de ser invariantes. Las preferencias son cambiantes, las posibilidades de acción alternativas y las relaciones de poder dominantes. Las innumerables imperfecciones y la incertidumbre en los datos de un sistema matemático de planificación, junto con la relatividad misma de los presupuestos políticos, reducen el dominio de validez del adjetivo óptimo que acompaña a sus resultados. La exacta satisfacción de sus restricciones no puede considerarse un requisito inexcusable cuando no se conocen con demasiada precisión los parámetros de sus ecuaciones representativas.

En cambio, la información dual, referente a la sensibilidad del objetivo frente a perturbaciones o distintas especificaciones del modelo pasa a ser el elemento fundamental e insustituible del análisis.

Fundamental porque facilita a los responsables políticos un medio de valorar las consecuencias de sus decisiones, testar la coherencia de las mismas con sus objetivos y analizar los compromisos que su actuación encierra.

Insustituible porque tal información no puede obtenerse de otra forma. La complejidad de un proceso dinámico escapa a cualquier tipo de análisis verbal. Bajo la hipótesis de que los responsables del sistema económico tratan de instrumentar racionalmente sus decisiones en la búsqueda de ciertos objetivos, (no necesariamente los declarados públicamente como tales), están actuando de acuerdo con un cierto modelo. Las más simples configuraciones mentales acerca de actos y consecuencias constituyen un modelo. Cuando los modelos quedan reducidos a una tan primitiva forma, poseen el reconfortante inconveniente de no ser explícitos, y de no mostrar claramente hasta donde llega el conocimiento responsable y donde empieza el riesgo calculado con el que nos enfrentamos a la incertidumbre.

El sistema dual de precios representado por $\psi(t)$ se deriva del proceso de optimización y está indisolublemente asociado con el plan óptimo resultante. En este sentido, el sistema de precios de una economía planificada es conceptualmente superior a los que produce y utiliza una economía de mercado. Considerar a estos como un sistema de valor común a la sociedad que los genera y utilizarlos como una medida homogénea y por tanto aditiva, es "identificar el funcionamiento real de la economía con el idealizado esquema walrasiano" (Albony (A7)). Las hipótesis simplificadoras que condicionan su validez se han olvidado o no se explicitan, cosa imposible de hacer con el modelo matemático que ha generado sus variables duales. La econo-

mía es sin duda el campo donde, como dice Albouy (A7), "l'ambiguité est le précédent le plus direct de l'imposture".

Pero cualquiera que sea la importancia que se acuerda conceder a la información dual y a sus aplicaciones, queda todavía por resolver el problema de su cálculo.

Este obstáculo, interpuesto entre la concepción teórica y la aplicación práctica, merece mucha más atención de la que con frecuencia recibe. A él nos referíamos implícitamente cuando recordábamos en la introducción que el principio del Máximo no es, a pesar de toda su importancia, un procedimiento de resolución de problemas reales. Solamente en algunos de ellos, como los lineal-cuadráticos (apéndice IV) conduce directamente a procedimientos standards de cálculo.

En la práctica, la primera de nuestras hipótesis es tan ilusoria como las otras, y la información dual no puede calcularse en todos los casos ni con toda la exactitud deseada. Señalemos, a título de ejemplo, que los algoritmos de penalización, en los que los multiplicadores se obtienen numéricamente como el producto de dos términos, uno de los cuales tiende a cero y el otro a infinito, generan la información con una indeseable imprecisión.

El estudio de los algoritmos de resolución de problemas como los planteados en los Apéndices I y II es un campo tan extenso como importante, que está dominado por las posibilidades que ofrece y las servidumbres que impone su gran protagonista: el ordenador digital.

A ellos nos referimos en el siguiente capítulo, anali-

zando las ventajas e inconvenientes de los distintos métodos en función del papel que en ellos juegue la dualidad. Un más profundo análisis de su influencia en el propio proceso de cálculo, que desborda su papel interpretativo de las soluciones, es considerado en el capítulo IV.

En el caso de juegos diferenciales, la matriz $\Psi(t)$ es muy difícil de obtener. Solamente el juego lineal cuadrático ha sido objeto de un tratamiento que conduce a algoritmos más o menos eficientes ($S3, S7$). En los demás casos, la búsqueda de soluciones, utiliza procedimientos de descomposición-coordinación ficticios que están actualmente en una fase incipiente de desarrollo (Pau ($P3$)).

METODOS DE SOLUCION.

- "If you give to the computer a bunch
of silly things, the only thing you
will get is a bunch of silly things...
A good theory is worth a hundred com
puter runs".

G.B. Dantzig.

III. 1 INTRODUCCION

III. 2 METODOS DIRECTOS Y PROGRAMACION MATEMATICA.

III. 3 METODO PRIMAL DE RESOLUCION: UN ALGORITMO DEL TIPO
GRADIENTE REDUCIDO GENERALIZADO.

III. 3. 1 INTRODUCCION

III. 3. 2 ESTRUCTURA BASICA DE LOS METODOS GRG.

III. 3. 3 APLICACION A PROBLEMAS DISCRETOS DE CONTROL
OPTIMO. CALCULO DEL GRADIENTE REDUCIDO.

III. 3. 4 CONCLUSIONES.

III. 4 METODOS DUALES DE RESOLUCION.

III. 4. 1 INTRODUCCION A LOS METODOS DUALES.

III. 4. 2 ALGORITMO BASADO EN EL METODO DUAL DE LOS
MULTIPLICADORES DE HESTENES.

III. 4. 3 ALGORITMO DUAL PURO TIPO LASDON-TAMURA.

Este capítulo es esencialmente de naturaleza operativa. En el exponemos tres procedimientos numéricos de resolución de problemas de control óptimo, basados en métodos de programación matemática.

En el capítulo I hemos señalado la importancia de la información dual (variables adjuntas y multiplicadores), en el análisis por el principio del máximo de problemas económicos de control óptimo.

Desarrollando esta idea, en el capítulo II hemos expuesto las distintas consideraciones conceptuales acerca de dichas variables y el uso de las mismas en un análisis económico.

Pero la utilización operativa de tales elementos exige su conocimiento previo. Toda su importancia conceptual se desvaloriza si no es posible obtener, en un problema determinado, el conjunto concreto de valores de $\psi(t)$ que permitan, en última instancia, el desarrollo real del análisis.

El esfuerzo para hacer corresponder valores a las magnitudes que instrumentan conceptos económicos no se identifica con el análisis de los problemas económicos. Pero sí es la condición previa e ineludible del mismo.

En consecuencia, la idea básica de los algoritmos desarrollados en este capítulo es la de conseguir una información dual fiable de una forma eficiente.

En el apartado III. 2 nos referimos brevemente a la disyuntiva entre métodos directos e indirectos en la resolución práctica de problemas de control óptimo.

Damos las principales referencias y consideraciones del proceso de discretización que permite convertirlos en pro-

blemas de programación matemática. Tal relación siendo de naturaleza más inmediata en problemas discretos de control, exponemos las razones que permiten plantear en tal forma determinados problemas de planificación económica. En contrapartida, nos referimos a la asociación de características de la naturaleza económica del problema y de su estructura matemática que inducen un planteamiento en tiempo continuo.

En la resolución de problemas de control óptimo discretos no lineales, es posible adoptar dos enfoques alternativos, que generan respectivamente las dos grandes familias de métodos primales y duales .

Ambos poseen, intrínsecamente asociados, una serie de ventajas e inconvenientes relativos. La exposición de un método primal en el punto III. 3 y el de dos métodos duales en el III. 4 nos permitirá analizar comparativamente tales diferencias de enfoque, sus hipótesis asociadas, exigencias de cálculo y los condicionamientos de su aplicación a los problemas objeto de nuestro interés.

Antes de entrar en este análisis detallado, recordemos aquí sus diferencias esenciales :

Los métodos primales no dependen, para su aplicación, de la convexidad del problema considerado. Mantienen la factibilidad de las sucesivas soluciones de su proceso iterativo a costa de un proceso de cálculo especialmente complejo numéricamente que precisa además el conocimiento de una solución inicial factible. La información de tipo dual generada en el propio proceso de cálculo que pueda utilizarse en un análisis de sensibilidad de las soluciones es, en general, pobre.

Los métodos duales precisan que el problema reúna ciertas condiciones de convexidad para asegurar su eficaz aplicación. Puesto que no se preocupan de garantizar la factibilidad de las sucesivas soluciones, poseen un proceso de cálculo menos complejo y no precisan el conocimiento de una solución inicial factible. La información que generan para efectuar un análisis de sensibilidad es especialmente rica.

Con estas características generales, aparece evidente que las líneas de evolución positiva se sitúan en la elaboración de :

- a) métodos primales que generen información dual como parte de su propio proceso de cálculo.
- b) métodos duales que amplíen la gama de problemas a los que pueden ser aplicados y/o que permitan resolver problemas estrechamente relacionados con el propuesto.

De acuerdo con ello, el método primal basado en el método del gradiente reducido generalizado expuesto en III. 3 efectúa el cálculo de las variables adjuntas como parte del mismo proceso de optimización.

A su vez, los dos algoritmos de III. 4 exponen como puede ser abordada la resolución de problemas no convexos mediante métodos duales. En el segundo de ellos se supone verificada la hipótesis de convexidad local mientras que en el primero se utiliza la transformación de Hestenes en problemas que no verifican siquiera tal hipótesis.

El interés de su presentación conjunta es el de observar las beneficiosas consecuencias derivadas de la convexidad local del problema. El algoritmo que utiliza dicha hipótesis permite la descomposición del primer nivel en sub-problemas, y

la consideración de efectos retardados en las ecuaciones de estado.

Los procesos de descomposición numérica que permiten instrumentar los métodos duales en problemas con Lagrangiano descomponible aditivamente, serán desarrollados en los sistemas de control multiniveles del capítulo V. Allí trataremos dos puntos fundamentales, relacionados con las consideraciones básicas aquí expuestas:

- las condiciones que deben reunirse para que los métodos duales de descomposición sean aplicables a problemas no convexos.
- la utilización de los teoremas de Everett para utilizar positivamente la información obtenida en un intento de resolución del problema primal propuesto, mediante su inserción en una familia de problemas del mismo tipo con segundos miembros de las restricciones parametrizados.

Todo el presente capítulo se basa en métodos de programación matemática. En el Apéndice III se ofrece un resumen estructurado del desarrollo de tales métodos, en cuyo contexto se encuadran los aquí utilizados.

Tanto la teoría de control óptimo como la programación matemática pueden ser divididas en dos grandes partes: el estudio de las propiedades que deben satisfacer las trayectorias o puntos solución y el desarrollo de métodos de cálculo que permitan obtener tales soluciones.

El principio del máximo, como las condiciones de Kuhn y Tucker, pertenecen a la primera de estas partes. Ambas son condiciones necesarias que deben satisfacer las soluciones pero no son, en sí mismos, un procedimiento práctico para obtenerlas.

Los métodos de cálculo de las soluciones pueden ser clasificados en dos grandes grupos.

Entendemos por métodos indirectos aquellos que parten del conjunto de condiciones necesarias enunciadas por el principio del máximo para determinar las trayectorias que las satisfacen. Generalmente ello se efectúa a través de un proceso iterativo a partir de una solución nominal que satisface previamente alguna de tales condiciones.

En cualquier caso, se observará que las condiciones necesarias se enuncian a través de un TPBVP y que la satisfacción de las mismas exige su resolución. Numerosos procedimientos han sido diseñados para conseguirlos con mejor o peor fortuna. Se trata de un complejo problema numérico que se complica adicionalmente con la existencia de restricciones sobre la trayectoria. Una clasificación de la ingente bibliografía al respecto puede encontrarse en Polak (p10).

Los métodos directos son aquellos que tratan de determinar, mediante un proceso iterativo de búsqueda inteligente a través del espacio de soluciones, aquellos elementos del mismo que satisfagan las condiciones de optimalidad. Estás

'son pues las que proporcionan el armazón conceptual de ambos tipos de procedimientos, pero de alguna manera podríamos decir que en los primeros son utilizadas " a priori " y en los segundos " a posteriori ".

Lo mismo ocurre en programación matemática, donde los métodos indirectos no se revelan de utilidad práctica directa sino en problemas de muy baja dimensionalidad.

Cualquiera que sea la formulación de un problema de control óptimo, el necesario uso del ordenador en su resolución exige que una u otra forma de discretización se lleve a cabo durante el proceso de cálculo. En los procedimientos indirectos la resolución del correspondiente TPBVP precisa la integración numérica de un sistema de ecuaciones diferenciales. La formulación continua del problema no debe hacer nos olvidar que el instrumento de cálculo lo discretizará ineludiblemente para poderlo resolver.

No es preciso insistir en la conocida relación entre los problemas de control óptimo y de programación matemática. El que el primero se reduce a uno de los segundos, de dimensión finita o infinita, según que su formulación sea discreta o continua, se deriva de la inmediata constatación de que las únicas variables que gobiernan la trayectoria, luego el valor de la funcional objetivo, son $u(t)$, $t \in [t_0, T]$, $x(t_0)$, con las ecuaciones de estado representando un conjunto de restricciones igualdad. Una muy completa exposición de las relaciones entre ambos campos puede encontrarse en Luenberger (L11).

Los métodos directos tratan de aplicar determinados algoritmos de programación matemática al problema de control óptimo. En los problemas continuos tal aplicación puede efectuarse de distintas formas. Se puede mantener la formulación continua del problema y aplicar las técnicas de pro-

gramación matemática extensibles a espacios de dimensión infinita. Alternativamente, se puede discretizar, "a priori", convirtiendo el problema en uno de control óptimo discreto y aplicarle todas las técnicas de programación matemática en problemas de dimensión finita.

En este segundo caso el problema inmediato es el estudio de la convergencia de tales aproximaciones a la solución del problema original a medida que disminuye la amplitud de la discretización. Existen al respecto los resultados debidos a Cullum (C11), Kirillova (K9) y Daniel (D2).

Con referencia a nuestro interés más inmediato, tales resultados establecen que la solución por discretización de un problema de control de tiempo mínimo difícilmente convergerá a la solución del problema original. La discretización debiera efectuarse previa conversión del problema de tiempo libre en uno de tiempo final fijado con restricciones en el estado final.

A partir de un problema continuo discretizado, o directamente de un problema formulado en tiempo discreto, puede tratarse de resolverse el correspondiente programa matemático con tantas variables como resulten del producto del número de instantes de tiempo considerados por el número de componentes de los vectores de estado y de control.

La aplicación de las condiciones de Kuhn y Tucker no conducirá a otra cosa que al principio del máximo discreto (Apéndice II). Pero nada impide tratar de aplicar alguno de los métodos directos de programación matemática recientemente desarrollados (Apéndice III).

El inconveniente evidente de tal procedimiento es la elevada dimensionalidad del problema. La única esperanza de resolverlo radica en la adaptación del método a la especial estructura dinámica del problema, manifestada especialmente

a través de la estructura del jacobiano de las restricciones igualdad. De hecho, tal estructura es la responsable de la sencilla expresión del gradiente en un problema de control y de la eficacia de tal método.

Tres obras básicas han sido dedicadas a este enfoque operativo de los problemas del control óptimo mediante procedimientos de programación matemática (Canon, Cullum, Polak (C1), Polak (P9), Tabak y Kuo (T1)).

Tal conversión aproximada de un problema continuo de control óptimo del tipo formulado en el Apéndice I :

$$M_{\max} \left[\Phi(x(T)) + \int_{t_0}^T f_0(x(t), u(t), t) dt \right]$$

con:

$$\dot{x}(t) = f(x(t), u(t), t)$$

sujeto a:

$$g(x(t), u(t), t) = 0$$

$$q(x(t), u(t), t) \leq 0$$

donde:

$$x(t) \in E^n$$

$$u(t) \in E^m$$

$$E^n \times E^m \times E \xrightarrow{f} E^n$$

$$E^n \times E^m \times E \xrightarrow{f_0} E$$

$$E^n \times E^m \times E \xrightarrow{g} E^l$$

$$E^n \times E^m \times E \xrightarrow{q} E^{l'}$$

en un problema de control óptimo discreto del tipo formulado en el Apéndice II :

$$\text{Max}_{\mu(k)} \left[\sum_{k=0} f_0(x(k), \mu(k), k) + \Phi(x(N)) \right]$$

con:

$$x(k+1) = x(k) + f(x(k), \mu(k), k) \quad (k=0, N-1)$$

sujeto a:

$$g(x(k), \mu(k), k) = 0$$

$$q(x(k), \mu(k), k) \leq 0 \quad (k=0, N)$$

para asimilarlo al programa matemático:

$$\text{Max } F_0(z)$$

$$r(z) = 0 \quad f(z) \leq 0$$

puede efectuarse de dos formas:

a)

$$z = (x(0), x(1), \dots, x(N), \mu(0), \mu(1), \dots, \mu(N-1))$$

$$F_0(z) = \sum_{k=0}^{N-1} f_0(x(k), \mu(k), k) + \Phi(x(N))$$

$$r(z) \equiv \begin{bmatrix} f(x(0), \mu(0), 0) + x(0) - x(1) \\ \vdots \\ f(x(k), \mu(k), k) + x(k) - x(k+1) \\ \vdots \\ f(x(N-1), \mu(N-1), N-1) + x(N-1) - x(N) \\ g(x(0), \mu(0), 0) \\ \vdots \\ g(x(N), \mu(N), N) \end{bmatrix} = 0$$

$$f(z) \equiv \begin{bmatrix} q(x(0), \mu(0), 0) \\ \vdots \\ q(x(N)) \end{bmatrix} \leq 0$$

Tal formulación tiene los inconvenientes asociados con la elevada dimensionalidad del vector \mathbf{z} .

$$b) \quad \mathbf{z} = (x(0), u(0), \dots, u(N-1))$$

El vector de estado en cada instante $x(k)$ queda determinado a partir del estado inicial $x(0)$ y la secuencia de controles aplicados $u_k = \{u(0), \dots, u(k-1)\}$ a través de la aplicación reiterada de las ecuaciones de estado.

Explicitamente :

$$x(k) = \varphi_k(x(0), u_k)$$

$$F_0(z) = \sum_{k=0}^{N-1} f_0(\varphi_k(x(0), u_k), u(k), k) + \Phi(\varphi_N(x(0), u_N))$$

$$r(z) \equiv \begin{bmatrix} g(x(0), u(0), 0) \\ \vdots \\ g(\varphi_N(x(0), u_N), N) \end{bmatrix} = 0$$

$$f(z) \equiv \begin{bmatrix} q(x(0), u(0), 0) \\ \vdots \\ q(\varphi_N(x(0), u_N), N) \end{bmatrix} \leq 0$$

En esta segunda formulación se reduce la dimensionalidad del vector \mathbf{z} y desaparecen explícitamente las restricciones igualdad correspondientes a las ecuaciones de estado. Pero el número de cálculos previos necesarios a la expresión de las funciones φ_k dificulta la aplica-

ción salvo en los casos en los que se conozca la forma de la solución general de las ecuaciones de estado. Tal es el caso cuando el sistema posea una dinámica lineal.

En los campos donde nacieron y se desarrollaron los modelos y técnicas de control óptimo, los fenómenos estudiados tenían una clara naturaleza continua. Los modelos discretos o las discretizaciones de modelos continuos no eran sino el reconocimiento explícito de una limitación de medida o una aproximación de cálculo.

Evidentemente la actividad económica tiene lugar de una forma continua. Pero cabe preguntarse acerca de la conveniencia de una modelización en tiempo continuo del proceso de control de sistemas económicos. La teoría del crecimiento óptimo utiliza una formulación continua en sus modelos. Pero ya hemos indicado que tal teoría no tiene ninguna pretensión de operatividad. Por el contrario, una serie de razones indican que una formulación discreta es mas adecuada:

- las constantes de tiempo características del proceso (la "rapidez" de su dinámica) hacen que se pueda plantear modelos representativos con intervalos de discretización lo bastante grandes para mantener su operatividad.
- los controles pueden identificarse facilmente con las funciones constantes por intervalos con un número finito de discontinuidades que exige la formulación discreta o la solución por discretización. Muchas de las variables de decisión económicas sólo pueden ser variadas a intervalos finitos de significativa amplitud.
- por razones políticas, institucionales y de su propia naturaleza los sistemas económicos son ejemplos típicos de aquellos en los que un exceso de variabili-

dad en los controles bloquean su funcionamiento. Una formulación continua representaría un derroche superfluo de medios de cálculo.

- la forma en la que se tienen disponibles los datos, permite alimentar de manera mas natural e inmediata una formulación discreta.

El reconocimiento de tales hechos no debe hacer olvidar las ventajas de una formulación continua. Estas pueden evidenciarse por la intencionalidad perseguida con el modelo y/o por sus características matemáticas.

Cuando se pretenda estudiar las distintas opciones de desarrollo a muy largo plazo y sus características cualitativas, los fenómenos de acumulación y la forma de las trayectorias se analizan mejor utilizando un lenguaje de ritmos que basándose en uno de saltos discretos. Cuando la variable sobre la que se define el óptimo es la propia variable temporal, (problemas de mínima duración de transición) una formulación continua es mas apropiada.

En cualquiera de ambos casos lo que se pretende en última instancia es estudiar las alternativas de realización de cambios estructurales mediante un modelo mas idealizado en su devenir temporal.

Si a ello se suma la consideración de un sistema con dinámica lineal, la posibilidad de encontrar soluciones analíticas al sistema de ecuaciones diferenciales lineales que lo representa, y por lo tanto de utilizar una formulación como la segunda de las indicadas, hace mas elegante y práctica una formulación continua.

Dado el interés de considerar relaciones no lineales, tanto en objetivos como ecuaciones de estado, los dos puntos siguientes se refieren a sistemas discretos no lineales.

III. 3.1 INTRODUCCION.

Entre los métodos de programación matemática, los algoritmos del tipo Gradiente Reducido Generalizado (GRG), han demostrado ser especialmente eficaces.

Como indica el Apéndice III, son del tipo primal, gozando como tales de las ventajas e inconvenientes allí reseñados.

La denominación Gradiente Reducido es debida a Wolfe (w2), quien desarrolló el método para el caso de restricciones lineales. El adjetivo Generalizado fue introducido por Abadie y Carpentier (A2) al extender el método de Wolfe al caso de restricciones no lineales.

Sus características han sido estudiadas teóricamente por Faure y Huard (F1) y su convergencia relacionada con la del Gradiente Projectado de Rosen por Luenberger (L13).

Todos los algoritmos GRG presentan una estructura general común . Esta consiste en la división de las variables del problema en un conjunto de variables básicas y un conjunto de variables no básicas, hecha posible por la aceptación de una hipótesis de no degeneración que no es sino la consecuencia de la satisfacción de los requisitos necesarios para la aplicación de las condiciones de Kuhn y Tucker ("constraints qualification").

De acuerdo con su estructura básica común, todos los algoritmos GRG convierten el problema original en un problema

reducido cuyas únicas restricciones son los límites individuales de variación de las variables no básicas.

En general, todos ellos pasan por el cálculo del Gradiente Reducido, cuyas componentes no son sino los multiplicadores de Kuhn y Tucker asociados a las restricciones del problema reducido. Sus principales diferencias son debidas a :

- El proceso de cálculo del gradiente reducido.

Ello depende de la particular estructura del problema. Este es un punto de importancia capital cuando se trate de problemas dotados de una estructura muy definida como son los problemas discretos de control óptimo.

- La utilización del Gradiente Reducido en el proceso de optimización.

Es decir, la definición de las direcciones de búsqueda en el problema reducido. De ello depende la eficacia global del método.

En el punto siguiente desarrollamos brevemente estas características de los algoritmos GRG. El lector que lo desee puede encontrar, en las referencias citadas, una más completa exposición del tema.

Apoyándonos en tales características, pasaremos a analizar en detalle la forma de calcular eficientemente el Gradiente Reducido en un problema discreto de control óptimo. Finalmente, describimos la utilización de tal información para conseguir una convergencia acelerada del algoritmo GRG resultante.

Cualquier programa matemático puede expresarse en la forma general:

$$\begin{aligned} \text{Min } f(x) \\ g_i(x) &= 0 \quad i=1, m \quad (1) \\ I_i \leq x_i \leq S_i \quad i=1, n \quad (2) \quad (n > m) \end{aligned}$$

mediante la introducción de variables de holgura positivas para convertir en igualdades las restricciones desigualdad, junto con la definición entre $-\infty, +\infty$ de las cotas I_i, S_i en las variables x_i . El vector x contiene tanto a las variables naturales del problema como a las posibles variables de holgura.

Para la aplicación de los métodos GRG, es preciso que se verifique la siguiente hipótesis de no degeneración :

En cualquier punto factible, es posible descomponer el vector x en dos vectores x_B, x_{NB} , de dimensiones m y $n-m$ respectivamente, tales que :

- a) El vector x_B , llamado vector de variables básicas o dependientes, verifique:

$$I_B < x_B < S_B$$

donde I_B, S_B son los vectores resultantes de la correspondiente descomposición de los vectores I, S .

- b) La matriz de dimensiones $(m \times m)$

$$B = \nabla_{x_B} g = \begin{bmatrix} \nabla_{x_B} g_1 \\ \vdots \\ \nabla_{x_B} g_m \end{bmatrix}$$

llamada matriz de base, sea no singular.

Consideremos a continuación :

- a) las implicaciones de dicha hipótesis.
- b) la forma en la que permite desarrollar el proceso de optimización.

De este análisis obtendremos como respectivas conclusiones :

- a) la hipótesis es cierta si se cumplen los requisitos necesarios para la aplicación de las condiciones de Kuhn y Tucker. No constituye pues una condición especialmente exigente, al menos teóricamente.
- b) el método GRG no sería aplicable sin la satisfacción de dicha hipótesis.

En efecto:

- a) Para que las condiciones de Kuhn y Tucker sean aplicables a un punto solución, este debe de ser un punto regular de las restricciones que en él sean activas. (Luenberger (L13)). En otras palabras, los gradientes de todas las restricciones activas simultáneamente deben ser linealmente independientes.

Si, como ocurre habitualmente, esta condición se supone verificada, las dos partes de la hipótesis de no degeneración se satisfacen necesariamente. En efecto:

De las $2n+m$ restricciones del problema, m son del tipo igualdad y $2n$ del tipo desigualdad. Puesto que el espacio de las variables x es de dimensión n , sólo pueden existir, como máximo, n restricciones activas simultáneas en cada punto regular.

Puesto que las m restricciones igualdad son activas en todo punto, un máximo de $n-m$ restricciones, de entre las $2n$ del tipo $I \leq x \leq S$, pueden ser activas simulta-

neamente. Ello implica que por lo menos m componentes de cada vector x , no estan tomando sus valores límites I_i, J_i . La existencia de las m variables básicas x_B :

$$I_i < x_{B_i} < J_i \quad (i=1, m)$$

está pues garantizada.

Supongamos, sin perder generalidad, que las m variables básicas x_B , son las m primeras componentes de x .

Con tal convención, el Jacobiano de las restricciones activas, que son en número $RA \leq n$, adopta la forma :

$$\nabla_x \text{ Restricciones activas.} = \begin{bmatrix} M_1 & \vdots & M_2 \\ \vdots & \ddots & \vdots \\ 0 & \vdots & \begin{matrix} 1 & & \\ & \ddots & \\ & & 1 \end{matrix} \end{bmatrix}$$

$\xleftarrow{\text{VNL}} \quad \xleftarrow{\text{VL}} \quad \xrightarrow{\text{restric desigualdad activas } (n-m)}$
 $\xrightarrow{\text{restric igualdad } (m)}$

donde :

$$M_1 = \nabla_{x_{NL}} g(x) \quad , \text{ de dimensiones } (m \times VNL \geq m)$$

$$M_2 = \nabla_{x_L} g(x) \quad , \dots \dots \dots (m \times (n-VNL))$$

x_{NL} = conjunto de variables que no toman valores límites.

x_L = conjunto de variables que estan tomando sus valores límites.

VNL = número de elementos del conjunto x_{NL} . Este número debe de ser, por lo menos, igual a m .

VL = número de elementos del conjunto x_L .

$$VL = n - VNL \leq n - m$$

Por la regularidad del punto x , todas las $(m + (n - VNL)) \leq n$ filas de esta matriz deben ser linealmente independientes. Lo mismo debe ocurrir con $(m + (n - VNL))$ de entre sus n

... columnas.

Puesto que las últimas $n - VNL$ lo son, la matriz M_1 debe poseer m columnas linealmente independientes.

Luego es posible escoger, entre las $VNL \geq m$ variables que verifican :

$$I_i < x_i < S_i$$

un conjunto de m de ellas, (las m variables básicas x_B) que verifiquen la segunda parte de la hipótesis de no degeneración:

$$\nabla_{x_B} g(x) \text{ regular}$$

Las restantes $n - m$ variables, que pueden estar o no en sus límites, constituyen las variables no básicas x_{NB} .

- b) Por el teorema de la función implícita, la regularidad de $\nabla_{x_B} g(x)$ es condición suficiente para que las ecuaciones $g(x) \equiv g(x_B, x_{NB}) = 0$ tengan una solución de la forma $x_B = x_B(x_{NB})$ en un entorno de un punto x verificando las restricciones (1). ($g(x) = 0$)

Por consiguiente, dadas las variables no básicas (0 independientes) x_{NB} , las m ecuaciones $g(x_B(x_{NB}), x_{NB}) = 0$ permiten determinar los valores de las variables básicas (0 dependientes) x_B . Estas pueden ser eliminadas de la función objetivo substituyéndolas por su expresión en función de las variables no básicas. El problema se expresa exclusivamente en función de estas últimas, en la forma :

$$\text{Min } f(x_B(x_{NB}), x_{NB}) = \text{Min } F(x_{NB})$$

$$I_{NB} \leq x_{NB} \leq S_{NB} \quad (3)$$

Este es el enunciado del problema reducido. Obsérvese que su planteamiento no hubiera sido posible sin la 2ª parte de

la hipótesis de no degeneración.
 La primera parte de dicha hipótesis permite abordar su proceso de solución. En efecto:

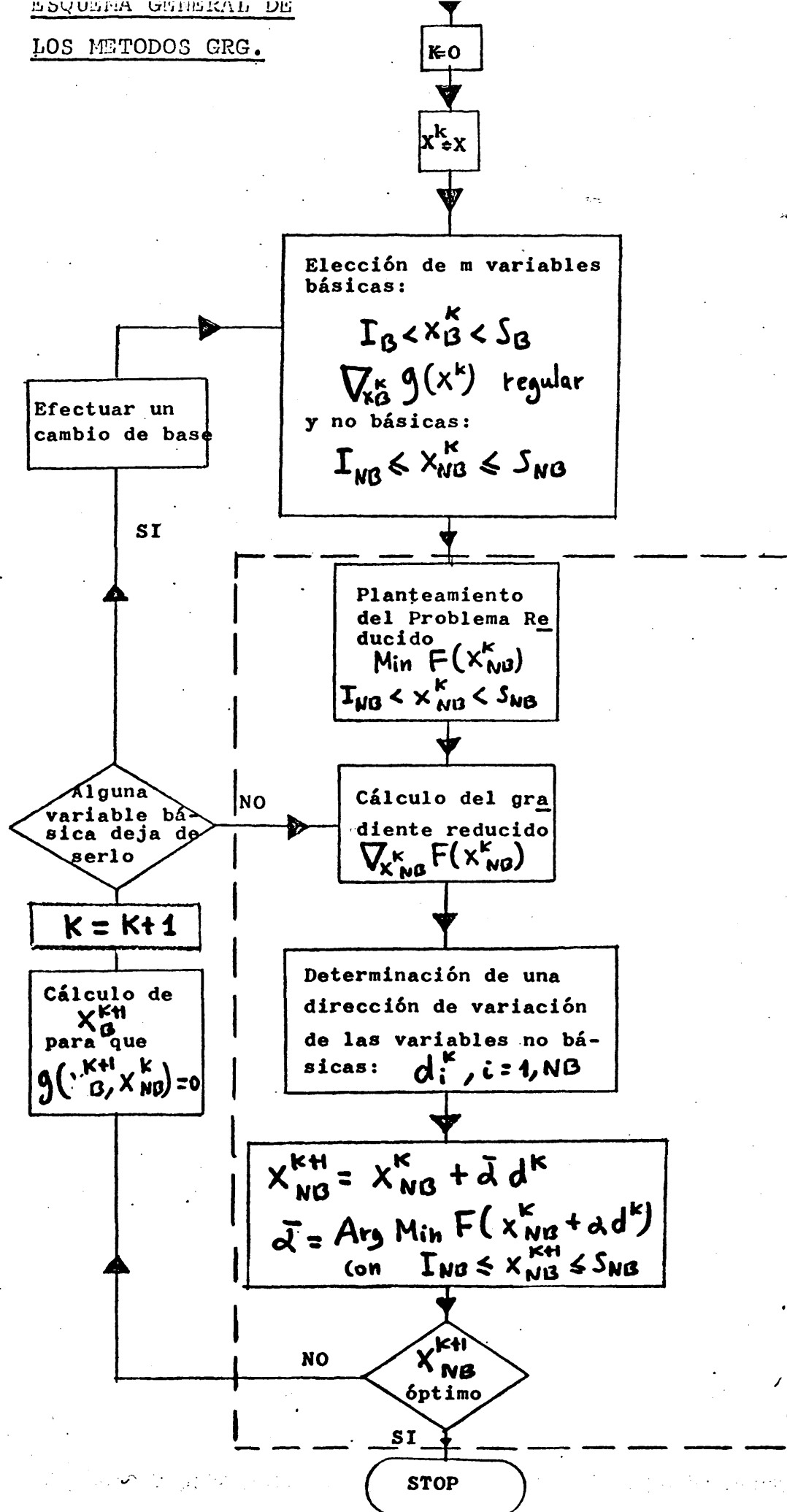
Para resolver el problema reducido, el algoritmo utilizado efectuará desplazamientos del punto factible X_{NB} , mediante un procedimiento que garantice la satisfacción de las restricciones explícitas (3). Pero cualquier modificación en las variables X_{NB} inducirá los necesarios cambios en las X_B para que las restricciones igualdad $g(X_B, X_{NB})=0$ se sigan verificando. Por ello, variaciones en X_{NB} , por pequeñas que fuesen, podrían provocar la violación de las restricciones $I_B < X_B < J_B$ (no explícitas en el problema reducido), si alguna de las X_B estuviera situada en los límites de su intervalo de variación en el punto X . Afortunadamente, la primera parte de la hipótesis de no degeneración asegura que todas las variables básicas verifican

$$I_B < X_B < J_B$$

A lo largo de las sucesivas iteraciones necesarias a la resolución del problema reducido, se modificará la composición del conjunto de variables básicas. El esquema general del método es el indicado en la figura FIII.1, donde aparece, encerrada en trazos, la parte correspondiente al problema reducido.

La coherencia teórica del método, no debe hacernos olvidar los múltiples problemas numéricos que encierra.

En primer lugar, salvo en los casos de restricciones lineales y otros de especial sencillez, no será posible obtener una forma analítica de la función $X_B(X_{NB})$. Normalmente habrá que recurrir a procedimientos iterativos para determinar numéricamente los valores de X_B correspondientes a específicos valores de X_{NB} .



tuarse analíticamente, sino que exigirá el cálculo de $F(x_{NB}^k + \alpha d)$ para distintos valores de α dentro de un procedimiento de búsqueda unidimensional. Para cada uno de estos valores será preciso evaluar los correspondientes valores de x_B .

Finalmente, el concepto numérico de regularidad de una matriz, presenta una ambigüedad que contrasta con su definición teórica. La matriz de base B puede estar lo bastante cerca de las condiciones de singularidad como para que su inversión plantee problemas de estabilidad numérica que deben ser objeto de especial atención en el algoritmo que desarrolle los cálculos en ordenador.

Enfrentado con estos problemas prácticos, la eficacia de un algoritmo particular dependerá del volumen de cálculo que realice por iteración y del número de estas que precise efectuar.

Lo primero depende de la forma en que se efectue el cálculo del gradiente reducido y el de las variables básicas. Lo segundo, de la eficacia de las direcciones de búsqueda d^k , definidas a partir del gradiente reducido.

Antes de pasar a analizar en detalle el primer punto, en el contexto particular de los problemas de control, centremos la atención en la expresión general del gradiente reducido, su significado en términos de las variables duales y en las condiciones del test de optimalidad del punto solución del problema reducido.

El gradiente reducido es el vector fila de $n-m$ compo-

hentes :

$$r = \nabla_{x_{NB}} F(x_{NB})$$

con:

$$r_i = \frac{\partial F(x_{NB})}{\partial x_{NB_i}} \quad (i=1, n-m)$$

Derivando $F(x_{NB})$ con respecto a cada una de las variables x_{NB_i} :

$$\frac{\partial F(x_{NB})}{\partial x_{NB_i}} = \frac{\partial f(x_{NB}, x_B)}{\partial x_{NB_i}} + \sum_{j=1}^m \frac{\partial f(x_B, x_{NB})}{\partial x_{B_j}} \frac{\partial x_{B_j}}{\partial x_{NB_i}}$$

es decir:

$$r_i = \frac{\partial f(x_{NB}, x_B)}{\partial x_{NB_i}} + \nabla_{x_B} f(x_B, x_{NB}) \frac{\partial x_B}{\partial x_{NB_i}}$$

con lo que el vector r viene dado por la expresión :

$$r = \underbrace{\nabla_{x_{NB}} f(x_B, x_{NB})}_{(1 \times n-m)} + \underbrace{\nabla_{x_B} f(x_B, x_{NB})}_{(1 \times m)} \underbrace{\nabla_{x_{NB}} x_B}_{(m \times n-m)}$$

en la que :

$$\nabla_{x_{NB}} x_B = \left[\frac{\partial x_{B_i}}{\partial x_{NB_j}} \right] = \begin{bmatrix} \nabla_{x_{NB}} x_{B_1} \\ \vdots \\ \nabla_{x_{NB}} x_{B_m} \end{bmatrix}$$

Esta última matriz puede expresarse a través de las restricciones $g(x_B, x_{NB}) = 0$, en la forma :

$$dg = \nabla_{x_B} g dx_B + \nabla_{x_{NB}} g dx_{NB} =$$

$$= \nabla_{x_B} g \nabla_{x_{NB}} x_B dx_{NB} + \nabla_{x_{NB}} g dx_{NB} = 0$$

$$\nabla_{x_{NB}} x_B = - [\nabla_{x_B} g]^{-1} \nabla_{x_{NB}} g$$

$(m \times n-m) \qquad (m \times m) \qquad (m \times n-m)$

Substituyendo esta expresión en la del gradiente reducido, esta queda :

$$r = \nabla_{x_{NB}} f(x_B, x_{NB}) - \nabla_{x_B} f(x_B, x_{NB}) [\nabla_{x_B} g]^{-1} \nabla_{x_{NB}} g$$

(4)

de donde cada uno de sus componentes viene dado por :

$$r_i = \frac{\partial f}{\partial x_{NB_i}} - \nabla_{x_B} f [\nabla_{x_B} g]^{-1} \frac{\partial g}{\partial x_{NB_i}}$$

(i=1, n-m)

El cálculo del gradiente reducido pasa pues por el de la inversa de la matriz de base, de dimensiones (m x m), cuya regularidad viene garantizada por la hipótesis de no degeneración.

Si el problema posee restricciones no lineales, la matriz de base debe ser calculada e invertida en cada iteración. Si las restricciones son lineales, la matriz de base es constante, lo cual exime de realizar un elevado volumen de cálculos. En tal caso, si las restricciones $g(x)=0$,

se escriben :

$$Ax - b = 0$$

y B, C son las matrices resultantes de efectuar la partición de A de acuerdo con la llevada a cabo en X entre variables básicas y no básicas:

$$[B \ C] \begin{bmatrix} x_B \\ x_{NB} \end{bmatrix} = b$$

al particularizar la expresión (4) del gradiente reducido, se obtiene la debida a Wolfe (110) :

$$r = \nabla_{x_{NB}} f - \nabla_{x_B} f B^{-1} C$$

El cálculo del gradiente reducido es un proceso laborioso, pero presenta la ventaja de facilitar al mismo tiempo los valores de todas las variables duales del problema. Dada la importancia que hemos asignado a la información dual, el obtenerla como una parte de su proceso es una de las grandes ventajas de los métodos GRG. Explicitemos la relación existente entre variables duales y gradiente reducido:

Las variables duales son los multiplicadores asociados con las restricciones igualdad (vector λ de m componentes) y los multiplicadores asociados con las restricciones desigualdad $I \leq x \leq S$ (vectores V_I, V_S , de n componentes)

Puesto que el par de restricciones $x_i \leq S_i, x_i \geq I_i$, correspondientes a una misma variable x_i , no pueden ser activas simultaneamente, los multiplicadores correspondientes V_{I_i}, V_{S_i} no pueden ser simultaneamente no nulos. Los vectores V_I, V_S podrían pues reducirse a un único vector, V . Manteniendo por el momento la distinción, el Lagrangiano del problema se escribe :

$$\mathcal{L} = f(x) + \lambda' g(x) + V_I' (I-x) + V_S' (x-s)$$

Las condiciones de optimalidad de Kuhn y Tucker, expresadas para los grupos de variables básicas y no básicas, se escriben :

$$\nabla_{x_B} f(x) + \lambda' \nabla_{x_B} g(x) - V_{I_B}' + V_{S_B}' = 0 \quad (5)$$

$$\nabla_{x_{NB}} f(x) + \lambda' \nabla_{x_{NB}} g(x) - V_{I_{NB}}' + V_{S_{NB}}' = 0 \quad (6)$$

$$V_I' (I-x) = 0$$

$$V_S' (x-s) = 0$$

$$V_S' \geq 0 \quad V_I' \geq 0$$

Por definición de variables básicas, los subvectores V_{S_B}' , V_{I_B}' son nulos. Gracias a ello, el valor de los multiplicadores asociados a las restricciones igualdad (λ) se obtiene inmediatamente de la ecuación (5):

$$\lambda' = - \nabla_{x_B} f(x) [\nabla_{x_B} g(x)]^{-1} \quad (7)$$

De acuerdo con (4), esta expresión es un paso necesario al cálculo del gradiente reducido. Como veremos después, es de la mayor importancia en problemas de control.

De acuerdo con (4) y (7), la expresión del gradiente reducido es :

$$\boxed{r = \nabla_{x_{NB}} f(x) + \lambda' \nabla_{x_{NB}} g(x)} \quad (8)$$

Substituyendo (8) en (6), se obtiene :

$$r - V'_{I_{NB}} + V'_{S_{NB}} = 0$$

$$r = V'_{I_{NB}} - V'_{S_{NB}}$$

Para cada componente i :

$$r_i = V_{I_{NB_i}} - V_{S_{NB_i}} \quad (i=1, n-m)$$

y puesto que cada variable X_{NB_i} puede encontrarse en una de las tres situaciones :

$$I_{NB_i} < X_{NB_i} < S_{NB_i} \implies V_{I_{NB_i}} = V_{S_{NB_i}} = 0$$

$$X_{NB_i} = I_{NB_i} \implies V_{S_{NB_i}} = 0$$

$$X_{NB_i} = S_{NB_i} \implies V_{I_{NB_i}} = 0$$

los correspondientes valores que pueden tomar las componentes del gradiente reducido son :

$$r_i = 0$$

$$r_i = V_{I_{NB_i}} > 0$$

$$r_i = -V_{S_{NB_i}} < 0$$

Es decir, las componentes del gradiente reducido son, en el óptimo, los multiplicadores de Kuhn y Tucker asociados a las cotas de las variables no básicas.

Aparte de la posibilidad que ello ofrece de disponer, a un coste suplementario nulo, de la información dual, tales relaciones constituyen el enunciado del test de optimalidad a efectuar en cada iteración del problema reducido. En ausencia de las restricciones $I_{NB} \leq X_{NB} \leq S_{NB}$, exigiríamos a un punto solución del problema reducido que su gradiente se anulara. Para tenerlas en cuenta exigimos que:

$$r_i > 0 \quad \text{si} \quad X_{NB_i} = I_{NB_i}$$

$$r_i < 0 \quad \text{si} \quad X_{NB_i} = S_{NB_i}$$

lo cual, como hemos visto, no hace sino traducir la satisfacción de las condiciones de Kuhn y Tucker en el problema original.

Evidentemente, si el problema planteado fuera uno de maximización, la expresión del gradiente reducido sería :

$$r = V'_{S_{NB}} - V'_{I_{NB}}$$

por lo que las condiciones de optimalidad del problema reducido serían:

$$r_i = 0 \quad \text{si} \quad I_{NB_i} < X_{NB_i} < S_{NB_i}$$

$$r_i > 0 \quad \text{si} \quad X_{NB_i} = S_{NB_i}$$

$$r_i < 0 \quad \text{si} \quad X_{NB_i} = I_{NB_i}$$

Una vez explicitados la estructura de los métodos GRG, el método general de cálculo del gradiente reducido, su significado en términos de las variables duales y las condiciones de optimalidad del problema reducido, pasemos a analizar su aplicación a un problema discreto de control óptimo.

III. 3. 3 APLICACION A PROBLEMAS DISCRETOS DE CONTROL OPTIMO. CALCULO DEL GRADIENTE REDUCIDO.

En un problema discreto de control óptimo con n variables de estado y extendido a $N+1$ instantes discretizados del tiempo, las ecuaciones de estado representan un conjunto de Nn restricciones igualdad.

Las restricciones desigualdad instantáneas $g(x(k), u(k), k) \leq 0$ explicitan las zonas del espacio de estado y de control por el que no se permite el discurrir de las trayectorias. En el más simple de los casos, y a la vez uno de los que más directamente traducen las exigencias reales, tales restricciones instantáneas adoptan la forma :

$$(1) \quad I E(k) \leq x(k) \leq S E(k) \quad k=0, \dots, N$$

$$(2) \quad I C(k) \leq u(k) \leq S C(k) \quad k=0, \dots, N-1$$

Añadiendo tales restricciones a la formulación de un problema discreto de control óptimo, como los expuestos en el Apéndice II :

$$Max J = \sum_{k=0}^{N-1} f_0(x(k), u(k), k) + \Phi(x(N))$$

$$x(k+1) = f(x(k), u(k), k) \quad k=0, \dots, N-1$$

$$x(k) \in E^n$$

$$u(k) \in E^m$$

su resolución numérica se complica notablemente, fundamen-

talmente debido a las restricciones (1).

De acuerdo con lo expuesto en III.2 , el problema anterior puede formularse como uno de los considerados en III. 3. 2, definiendo el nuevo vector de variables :

$$x' = [x(0), x(1), \dots, x(k), \dots, x(N), u(0), \dots, u(k), \dots, u(N-1)]$$

de dimensión $(N+1)n + Nm$, con el vector de restricciones igualdad :

$$g(x) \equiv \begin{bmatrix} f(x(0), u(0), 0) - x(1) \\ \vdots \\ f(x(N-1), u(N-1), N-1) - x(N) \end{bmatrix} = 0$$

de dimensiones Nn , los límites en las variables :

$$I \leq x \leq S$$

$$I' = [I_E(0), I_E(1), \dots, I_E(N), I_C(0), \dots, I_C(N-1)]$$

$$S' = [S_E(0), S_E(1), \dots, S_E(N), S_C(0), \dots, S_C(N-1)]$$

y con el objetivo $\text{Max } J(x)$.

La aplicación de un método GRG a este problema, implica la división del conjunto X de variables en un vector de variables básicas x_B de dimensión Nn y un vector de $Nm + n$ variables no básicas de forma que se satisfaga la hipótesis de no degeneración. La división del conjunto de variables entre básicas y no básicas no tiene teóricamente nada que ver con la división de los componentes de X en variables de estado y de control, aunque, como veremos, la identificación de ambas divisiones tendría importantes consecuencias sobre la eficacia del cálculo.

En la expresión general del gradiente reducido :

$$r = \frac{dJ(x)}{dx_{NB}} = \nabla_{x_{NB}} J(x) - \nabla_{x_B} J(x) [\nabla_{x_B} g(x)]^{-1} \nabla_{x_{NB}} g(x) \quad (1)$$

sabemos que el elemento $-\nabla_{x_B} J(x) [\nabla_{x_B} g(x)]^{-1}$ representa el vector de variables duales asociados a las restricciones igualdad $g(x) = 0$. Puesto que estas son las ecuaciones de estado del problema, dicho vector lo componen las variables adjuntas $\psi(k), (k=1, N)$.

El cálculo de r implica pues, necesariamente, el de las variables adjuntas, satisfaciendo así la deseada condición de eficacia en el cálculo de la información dual. Sin embargo, la consideración de las dimensiones de los elementos utilizados revela que es preciso efectuar, en cada iteración, la inversión de una matriz de dimensiones $(N_n \times N_n)$ y la solución de un sistema no lineal de N_n ecuaciones. La solución global de ambos problemas representa un volumen de cálculo que puede comprometer la viabilidad práctica del método.

Es preciso, tener en cuenta la especial estructura del sistema de ecuaciones $g(x) = 0$, y por lo tanto de la matriz $\nabla_{x_B} g(x)$, derivada del carácter dinámico del problema (Mekra (M5), Davis (M5), Abadie (A3)) para arbitrar procedimientos que permitan:

- descomponer el cálculo de las variables básicas en función de las no básicas, en una serie de problemas de menor talla.
- calcular las componentes del gradiente reducido sin precisar la inversión de una matriz $(N_n \times N_n)$

Para ello utilizamos la propiedad, antes enunciada, de que los componentes del gradiente reducido son los multiplicadores de Kuhn y Tucker asociados a las cotas sobre las variables no básicas.

Para facilitar el cálculo de las variables básicas dadas las no básicas, sería extremadamente favorable que el vector X_B fuera el conjunto de las N_n variables de estado $X(1), \dots, X(N)$. De ser ello posible, el cálculo del vector X_B dado el vector :

$$X'_{NB} = [X(0), \mu(0), \mu(1), \dots, \mu(K), \dots, \mu(N-1)]$$

(las variables no básicas serían el estado inicial y la secuencia de controles), se limita a la solución secuencial de los N grupos (que no sistemas) de las n ecuaciones de estado :

$$X(K+1) = f(X(K), \mu(K), K) \quad (K=0, N-1)$$

Observe que ello implica suponer que las componentes del vector de estado no toman nunca sus valores límites. En tal caso, llamando :

$\Psi(K+1)$: vector de variables adjuntas, o variables duales asociados a:

$$X(K+1) = f(X(K), \mu(K), K)$$

$X(K)$: vector de variables duales asociados a :

$$I_C(K) \leq \mu(K) \leq S_C(K)$$

$$(K=0, N-1)$$

$\lambda(K)$: vector de variables duales asociados a :

$$I_E(K) \leq X(K) \leq S_E(K)$$

$$(K=0, N)$$

demostramos que el vector r , de expresión genérica (1),

está compuesto por $[\lambda(0), x(0), \dots, x(N-1)]$. Al así hacerlo, obtendremos un método de determinación secuencial de las variables adjuntas $\psi(k)$ y de las componentes de r .

De acuerdo con el Apéndice II, podemos escribir :

$$\nabla_{u(k)} f_0(x(k), u(k), k) + \psi'(k+1) \nabla_{u(k)} f(x(k), u(k), k) =$$

$$= x'(k) \quad (2)$$

$(k = 0, N-1)$

$$\nabla_{x(N)} \Phi(x(N)) - \psi'(N) = \lambda'(N) \quad (2')$$

$$\nabla_{x(k)} f_0(x(k), u(k), k) + \psi'(k+1) \nabla_{x(k)} f(x(k), u(k), k)$$

$$- \psi'(k) = \lambda'(k) \quad (3)$$

$(k = 0, N-1)$

En principio, observese que la expresión (3) adopta, para $k=0$, la misma forma genérica de (2), porque al no estar fijado el estado inicial, $\psi(0) = 0$. (Apéndice I).

Definamos el vector Λ :

$$\Lambda' = [\Lambda'(1), \Lambda'(2), \dots, \Lambda'(N)]$$

donde cada uno de los $\Lambda(k)$ es de dimensión $(n \times 1)$, como :

$$\Lambda' = -\nabla_{x_0} J(x) [\nabla_{x_0} g(x)]^{-1}$$

o, lo que es lo mismo :

$$\Lambda' \nabla_{x_B} g(x) = - \nabla_{x_B} J(x) \quad (4)$$

Teniendo en cuenta la definición de las restricciones $g(x)$, del objetivo $J(x)$ y de las variables básicas x_B (4) se desarrolla como sigue :

$$[\Lambda'(1), \dots, \Lambda'(N)] \begin{bmatrix} -I & \ddots & & & \\ \vdots & \ddots & \ddots & & \\ \nabla_{x(1)} f^{(1)} & -I & & & \\ \vdots & \ddots & \ddots & \ddots & \\ & \nabla_{x(2)} f^{(2)} & -I & & \\ \vdots & \ddots & \ddots & \ddots & \\ & & & \nabla_{x(N-1)} f^{(N-1)} & -I \end{bmatrix} =$$

$$= - \left[\nabla_{x(1)} f^{(1)}, \dots, \nabla_{x(N)} \Phi(x(N)) \right]$$

que se observa equivale al siguiente procedimiento recurrente para el cálculo de los subvectores $\Lambda'(k)$ a partir de $\Lambda(N)$:

$$\begin{aligned} \Lambda'(N) &= \nabla_{x(N)} \Phi(x(N)) \\ &\vdots \\ \Lambda'(k) &= \nabla_{x(k)} f^{(k)} + \Lambda'(k+1) \nabla_{x(k)} f^{(k)} \\ &\vdots \\ \Lambda'(1) &= \nabla_{x(1)} f^{(1)} + \Lambda'(2) \nabla_{x(1)} f^{(1)} \end{aligned} \quad (5)$$

que demuestra su coincidencia con las variables adjuntas del principio del máximo discreto definidas en el Apéndice II.

Luego el gradiente reducido, en el caso que las variables básicas fueran las variables de estado $x(k)$, $k=1, N$, se expresa como :

$$r = \nabla_{x_{NB}} J(x) + \psi' \nabla_{x_{NB}} g(x)$$

Más explícitamente :

$$\begin{aligned} & \left[\frac{dJ}{dx(0)}, \frac{dJ}{du(0)}, \dots, \frac{dJ}{du(N-1)} \right] = \\ & = \left[\nabla_{x(0)} f_0(0), \nabla_{u(0)} f_0(0), \dots, \nabla_{u(N-1)} f_0(N-1) \right] + \\ & + \left[\psi'(1), \dots, \psi'(N) \right] \begin{bmatrix} \nabla_{x(0)} f_0(0) & \nabla_{u(0)} f_0(0) & \dots & \dots & \dots \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \nabla_{u(1)} f_1(1) & \vdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \nabla_{u(N-1)} f_{N-1}(N-1) \end{bmatrix} \end{aligned}$$

Que, desarrollado, equivale a :

$$r_1 = \frac{dJ}{dX(0)} = \nabla_{X(0)} f_0(\mu(0), X(0), 0) + \Psi'(1) \nabla_{X(0)} f(X(0), \mu(0), 0)$$

$$r_2 = \frac{dJ}{d\mu(0)} = \nabla_{\mu(0)} f_0(\mu(0), X(0), 0) + \Psi'(1) \nabla_{\mu(0)} f(\mu(0), X(0), 0)$$

$$\vdots$$

$$r_{k+2} = \frac{dJ}{d\mu(k)} = \nabla_{\mu(k)} f_0(\mu(k), X(k), k) + \Psi'(k+1) \nabla_{\mu(k)} f(\mu(k), X(k), k)$$

$$\vdots$$

$$(6) \quad r_{N+1} = \frac{dJ}{d\mu(N-1)} = \nabla_{\mu(N-1)} f_0(\mu(N-1), X(N-1), N-1) + \Psi'(N) \nabla_{\mu(N-1)} f(\mu(N-1), X(N-1), N-1)$$

Es decir, los valores de $\lambda'(0)$ y de $X'(k)$, $k=0, \dots, N-1$ de acuerdo con las expresiones (3) para $k=0$ y (2) para $k=0, N-1$.

Con estos resultados, aparece evidente que la elección de las variables $X(k)$, $(k=1, N)$, como variables básicas, posible por hipótesis, se corresponde con la más eficiente forma de desarrollar las distintas partes del cálculo.

En efecto:

- La determinación de las variables básicas $X(k)$, $k=1, N$ dadas las no básicas, consiste en la aplicación reiterada a partir del instante inicial $(X(0), 0)$, de las ecuaciones de estado :

$$X(k+1) = f(X(k), \mu(k), k)$$

desde $k=0$ hasta $k=N-1$.

- Una vez conocidas las trayectorias de estado y de control $X(k), u(k)$, las variables adjuntas $\psi(k)$ pueden calcularse secuencialmente desde $K=N$ hasta $K=1$ utilizando las expresiones (5). Tales expresiones no son sino las ecuaciones (3) teniendo en cuenta que $\lambda'(k)=0, (k=1, N)$, puesto que las variables $x(1), \dots, x(N)$ son básicas.

Observese que ello equivale a calcular, secuencialmente, los valores de N grupos de n ecuaciones.

- Una vez calculadas las variables adjuntas, se calculan, una a una, las componentes del gradiente reducido mediante las ecuaciones (6).

Dotados de un procedimiento eficaz para calcular el gradiente reducido, construiremos, con base en este, unas direcciones de búsqueda que permiten encontrar una trayectoria para la que las componentes del gradiente reducido $(\lambda(0), \chi(0), \dots, \chi(N-1))$ satisfagan las condiciones de optimalidad del problema reducido:

$$\begin{array}{lll}
 r_{1i} = 0 & \text{si} & IE_i(0) < x_i(0) < SE_i(0) \\
 r_{1i} > 0 & \text{si} & x_i(0) = SE_i(0) \\
 r_{1i} < 0 & \text{si} & x_i(0) = IE_i(0) \quad (i=1, n) \\
 \\
 r_{ki} = 0 & \text{si} & IC_i(k-2) < \mu_i(k-2) < SC_i(k-2) \\
 r_{ki} > 0 & \text{si} & SC_i(k-2) = \mu_i(k-2) \\
 r_{ki} < 0 & \text{si} & IC_i(k-2) = \mu_i(k-2) \\
 & & (i=1, m)
 \end{array} \quad \left. \vphantom{\begin{array}{l} r_{1i} = 0 \\ r_{1i} > 0 \\ r_{1i} < 0 \\ r_{ki} = 0 \\ r_{ki} > 0 \\ r_{ki} < 0 \end{array}} \right\} \begin{array}{l} k=2 \\ \vdots \\ N+1 \end{array}$$

La dirección del propio gradiente reducido podría constituir una dirección de búsqueda. Dada la baja velocidad de convergencia del método del gradiente ello impli-

En la solución del problema reducido son decisivas las aportaciones de Abadie(A2) y Lasdon-Fox-Ratner(L2). Antes de referirnos brevemente a ellas en las conclusiones, consideremos ahora el caso en el que no todas las variables de estado puedan ser escogidas como básicas.

En realidad, aunque la adopción de tal base se corresponda con la más lógica y eficaz manera de escoger las variables dependientes, es difícil de mantener la hipótesis de que tal elección sea posible. El suponer que las restricciones instantáneas sobre el vector de estado permanecerán inactivas en todos los puntos de la trayectoria no puede corresponderse más que con una absoluta irrelevancia real de los factores que tales restricciones traducen. En general, pues, la trayectoria del vector de estado pasará por zonas de la frontera de valores admisibles en un intento de explotar al máximo las potencialidades del sistema dentro de las restricciones que le son definidas.

Cuando ello ocurra, los correspondientes componentes del vector de estado no podrán ser considerados como variables básicas. Deberán ser substituidos por otras variables, escogidas entre $X(0)$, $\mu(k)$, $(k=0, N-1)$, de forma que se verifique la hipótesis de no degeneración.

En consecuencia:

- 1) El vector de variables básicas X_B ya no estará formado por $X(k)$, $(k=1, N)$, sino por una mezcla de $X_i(k)$, $\mu_j(k)$ que variará para cada valor de k .
- 2) El gradiente reducido no estará compuesto por $\lambda(0)$, $X(k)$, $(k=0, N-1)$ sino por conjuntos de variables $\lambda_i(k)$, $X_j(k)$ correspondientes a los componentes

$x_i(k)$, $\mu_j(k)$ que no sean variables básicas.

- 3) Las condiciones del test de optimalidad del problema reducido deben ser expresadas con respecto a tales componentes de los vectores λ , x . Es decir, la trayectoria que se tome como solución debe verificar:

$\forall \mu_i(k)$ que sea no básica :

$(i=1, m)$
 $(k=0, N-1)$

$$x_i(k) = 0 \quad \text{si} \quad IC_i(k) < \mu_i(k) < SC_i(k)$$

$$x_i(k) \geq 0 \quad \text{si} \quad \mu_i(k) = SC_i(k)$$

$$x_i(k) < 0 \quad \text{si} \quad \mu_i(k) = IC_i(k)$$

$\forall x_i(k)$ que sea no básica :

$(i=1, n)$
 $(k=0, N)$

$$\lambda_i(k) = 0 \quad \text{si} \quad IE_i(k) < x_i(k) < SE_i(k)$$

$$\lambda_i(k) > 0 \quad \text{si} \quad x_i(k) = SE_i(k)$$

$$\lambda_i(k) < 0 \quad \text{si} \quad x_i(k) = IE_i(k)$$

Con tal composición del conjunto de variables básicas se habrá perdido, en general, la posibilidad de descomponer secuencialmente el cálculo de las variables básicas, de las adjuntas y del gradiente reducido.

Sin embargo, cabe analizar si tal imposibilidad es parcial o total.

Supongamos que cada una de las variables $x_i(k)$ que no

pueden ser tomadas como básicas por tomar valores límites en su intervalo de variación, puede ser substituida en la base por una variable de control $\mu_j(k-1)$ correspondiente al periodo anterior. En tal caso, la secuencialidad del cálculo puede ser mantenida, si bien cada una de sus etapas es ahora de mayor dificultad numérica. En efecto :

- 1) Para cada par de instantes consecutivos del tiempo, $K, K-1$, los vectores de estado $X(K)$ y de control $\mu(K-1)$ contienen n variables básicas. De ellas, $n - \pi(K)$ corresponden a componentes $x_i(K)$ y las $\pi(K)$ restantes corresponderían a componentes $\mu_j(K-1)$. En consecuencia, la obtención de las variables básicas dadas las independientes, consiste en la resolución secuencial de los N sistemas de n ecuaciones representados por las ecuaciones de estado, en el orden $K=1, N$.
- 2) El cálculo secuencial de las variables adjuntas desde $K=N$ hasta $K=1$ se consigue teniendo en cuenta que :

$$\forall \mu_i(K) \text{ que sea básica } \lambda_i(K) = 0$$

$$\forall x_i(K) \text{ que sea básica } \lambda_i(K) = 0$$

Por lo tanto, la ecuación (2), se escribe, $\forall x_i(N)$ que sea básica :

$$\psi'_i(N) = \frac{\partial \Phi(x(N))}{\partial x_i(N)}$$

lo que permite obtener inmediatamente $n - \pi(N)$ componentes del vector .

Los restantes $\pi(N)$ componentes se obtienen de la solución del sistema lineal de $\pi(N)$ ecuaciones obtenido expresando la ecuación (2), particularizada para $K=N-1$,

para cada uno de los componentes $\mu_j(N-1)$ que sean variables básicas:

$$\nabla_{\mu_j(N-1)} f_0(x(N-1), \mu(N-1), N-1) + \psi'(N) \nabla_{\mu_j(N-1)} f(x(N-1), \mu(N-1), N-1) = 0$$

Una vez calculado completamente el vector $\psi(N)$, es posible obtener secuencialmente los vectores $\psi(k)$, $(k=N-1, 1)$, y precisamente en este orden, por aplicación reiterada de los componentes de las ecuaciones (2) y (3) correspondientes a variables básicas. Observe-se que ello implica la solución sucesiva de $N-1$ sistemas lineales con $\pi(k)$ incógnitas.

- 3) Una vez calculadas las variables adjuntas $\psi(k)$, $(k=1, N)$, las ecuaciones (6) permiten calcular las componentes del gradiente reducido.

La utilización de bases constituidas por subvectores de la forma $(\mu_j(k-1), x_i(k))$, $(j=1, l \leq m, i=1, l' \leq n, l+l'=n, k=1, N)$ posee una ventaja adicional :

La matriz de base posee la estructura representada, a título de ejemplo, en la figura F III.2, por lo que su regularidad depende solamente de la regularidad de las sub-matrices diagonales. Cuando las variables básicas eran los vectores de estado $x(k)$, $(k=1, N)$, la regularidad estaba garantizada puesto que las submatrices diagonales eran las matrices unidad de dimensiones $(n \times n)$.

Del análisis de su estructura se desprende que la regularidad de cada una de las nuevas submatrices diagonales depende solamente de la matriz constituida por los elementos:

$$\frac{\partial f_i(x(k-1), \mu(k-1), k-1)}{\partial \mu_j(k-1)}$$

sea básica.

Ilustrándolo mediante un ejemplo, la regularidad de

-1				
-1	a		b	
	c	-1	d	
		-1		
			-1	
				-1

sólo depende de la regularidad de la matriz (2×2) formada por los elementos a, b, c, d :

a	b
c	d

Como puede deducirse por alternación de sus columnas. Ello facilita la selección de las nuevas variables básicas entre las $\mu_j(k-1)$ tales que :

$$IC_i(k-1) < \mu_j(k-1) < SC_i(k-1)$$

mediante un test de regularidad de la nueva matriz de base muy fácil de calcular.

Por supuesto, nada garantiza que se pueda encontrar un conjunto de componentes substitutorias $\mu_j(k-1)$ para cada uno de los instantes discretizados $k=1, N-1$. En otras palabras, nada garantiza que se puedan encontrar n variables dependientes en cada uno de los conjuntos $\mu(k-1); X(k)$, $\forall k=1, N$. En tal caso la recursividad total del cálculo no puede utilizarse. Sin embargo, entre las favorables

circunstancias antes descritas y la consideración global del jacobiano de las restricciones, es posible instrumentar soluciones intermedias.

Estas consisten en agrupar secuencias de períodos de tiempo, $k-1, k+l$, tales que el número de variables dependientes que contengan:

$$\mu_j(i-1), x_p(i) \quad \begin{matrix} i \in [k, k+l] \\ j \in [1, m], p \in [1, n] \end{matrix}$$

sea igual a n .

Considerando a tales agrupaciones de períodos consecutivos como un nuevo período simple, la secuencialidad del cálculo se mantiene sobre los nuevos períodos simples así definidos. Evidentemente la dimensión de los sistemas a resolver en cada uno de ellos es mayor que en el caso anteriormente considerado, en el que $l=0$. Pero así se evita al menos el plantear el problema en su dimensionalidad total.

El siguiente organigrama describe en su detalle este proceso de cálculo del gradiente reducido. Obsérvese que la composición del conjunto de variables básicas se modifica de iteración en iteración.

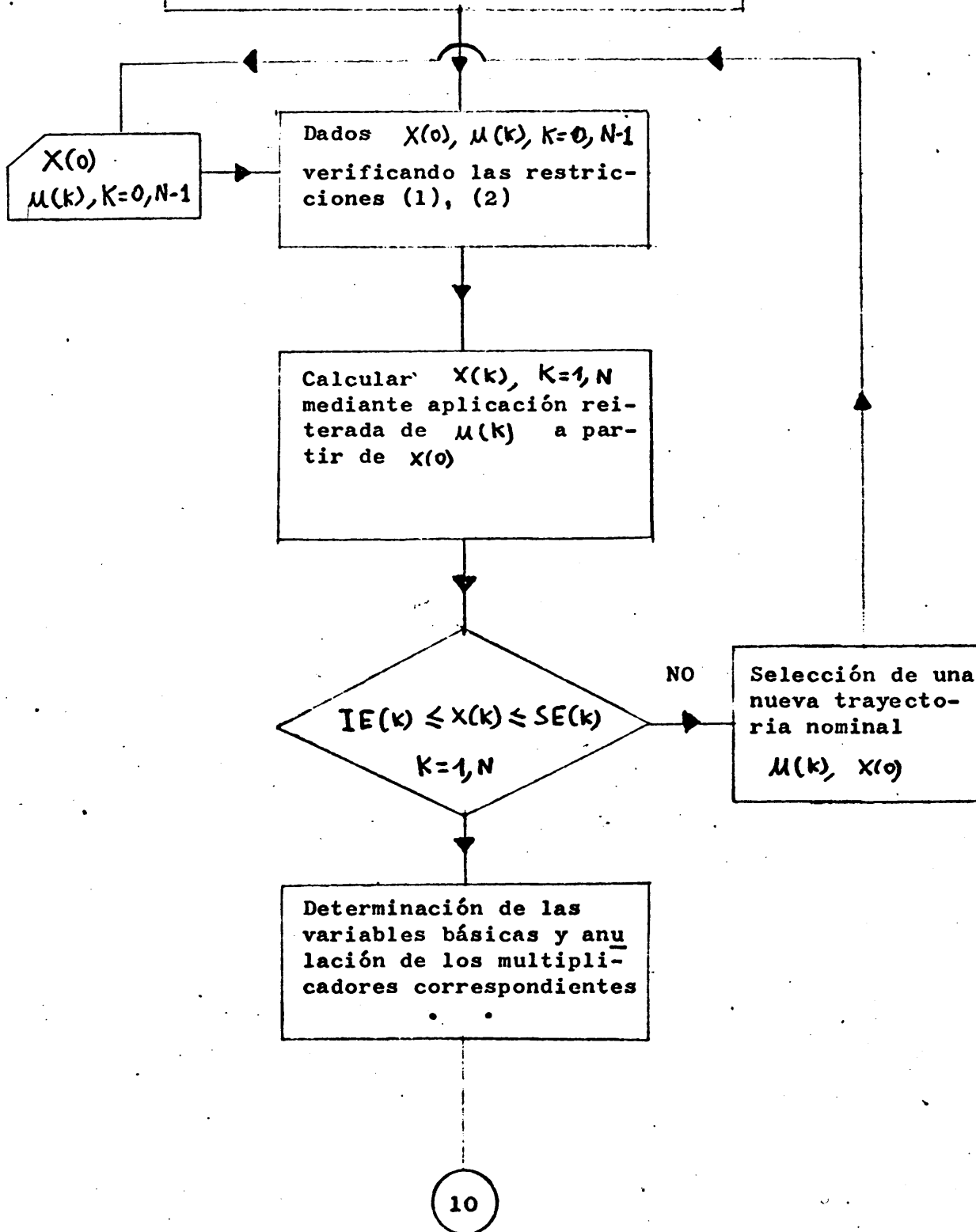
(Figura F III.3)

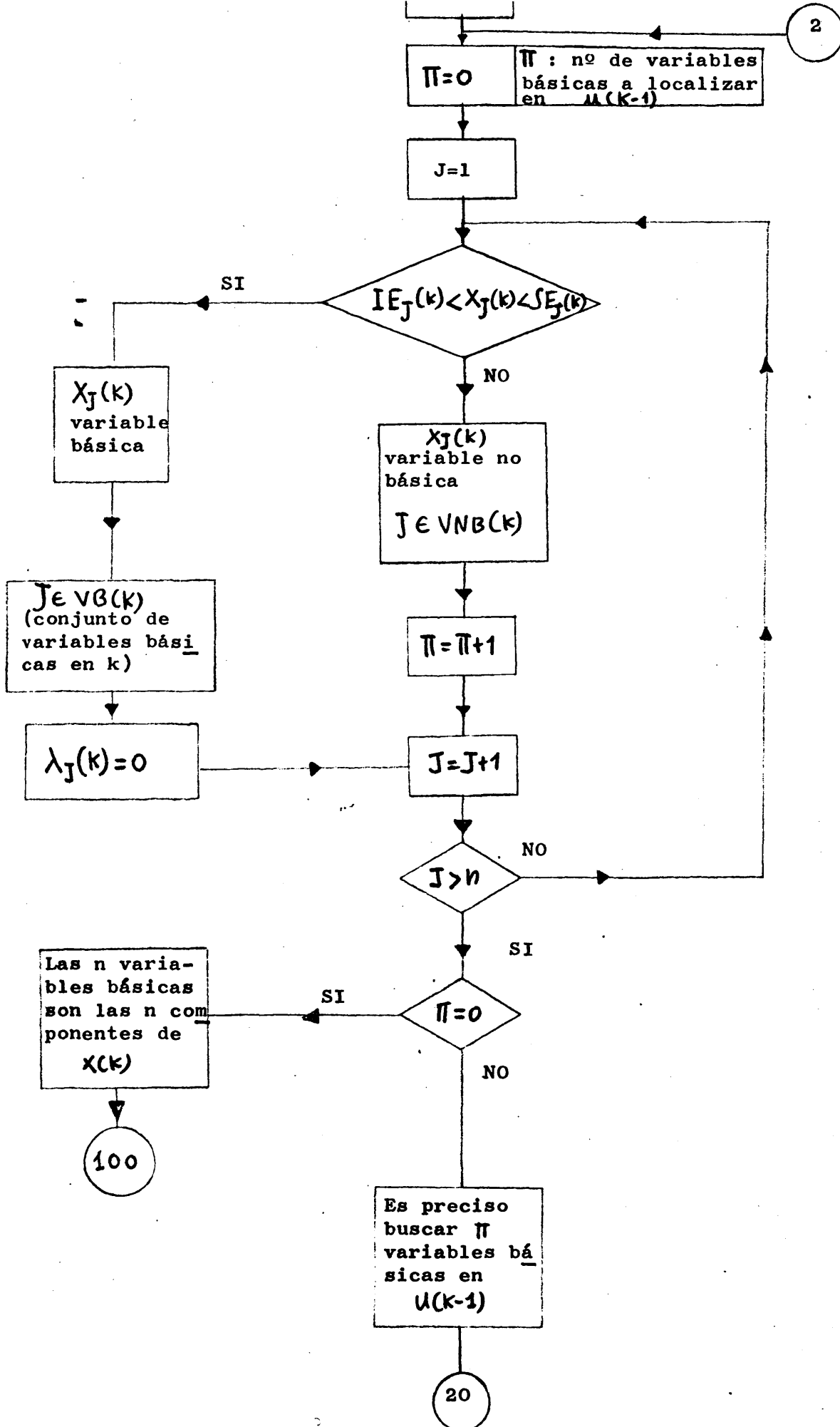
$$J = \Psi(x(N)) + \sum_{k=0}^{N-1} f_0(x(k), u(k), k)$$

$$x(k+1) = f(x(k), u(k), k) \quad (k=0, N-1)$$

$$IE(k) \leq x(k) \leq SE(k) \quad (k=0, N-1)$$

$$IC(k) \leq u(k) \leq SC(k) \quad (k=0, N)$$





NCB=0

I=1

NO

$$IC_I(k-1) < \mu_I(k-1) \leq k$$

SI

$\mu_I(k-1)$
no puede
ser bá-
sica.

$\mu_I(k-1) \in CB(k)$ CB(k): conjunto de com-
ponentes de $\mu(k-1)$ can-
didatas a básicas.

$$NCB = NCB + 1$$

Calcular

$$\nabla_{\mu_I(k-1)} f_J(x(k-1), \mu(k-1), k-1) \\ \forall J \in VNB(k)$$

I = I + 1

I > m

NO

SI

No hay bas-
tantes va-
riables can-
didatas a
básicas en
 $\mu(k-1)$ para
sustituir
a las $x_J(k)$
no básicas.

NO

$$NCB \geq \pi$$

SI

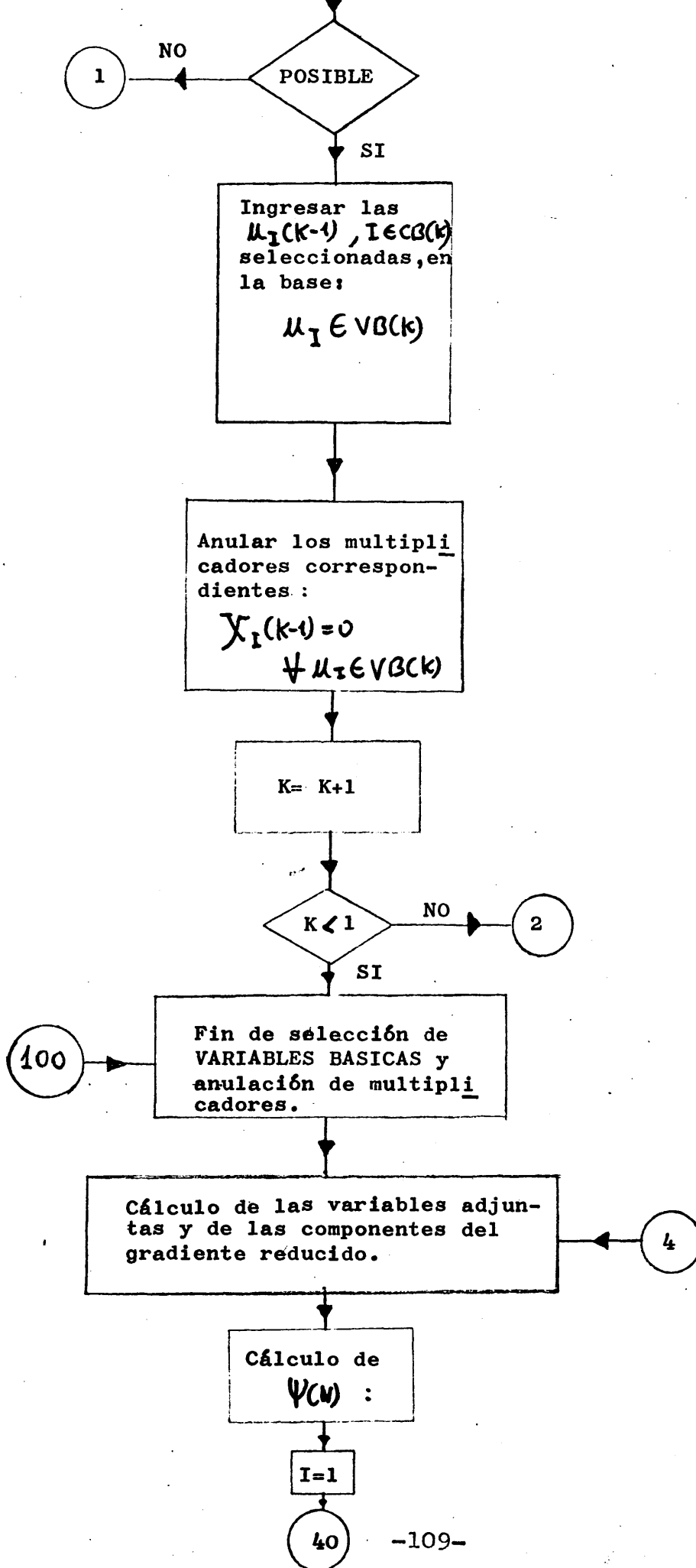
Seleccionar entre
las $\mu_J(k-1) \in CB(k)$
 π de ellas que
satisfagan

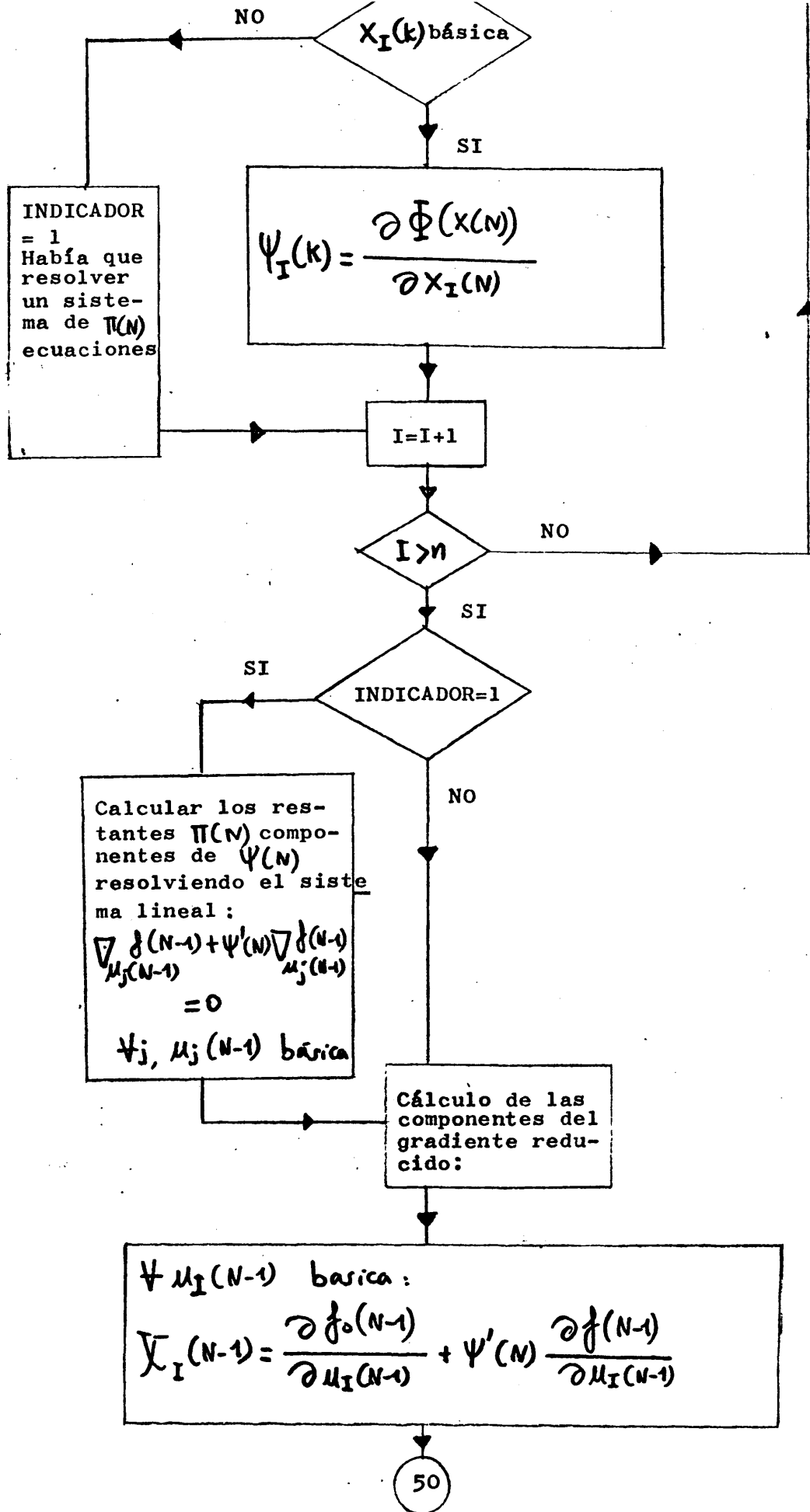
$$\det \left[\frac{\partial f_i(k-1)}{\partial \mu_j(k-1)} \right] \neq 0$$

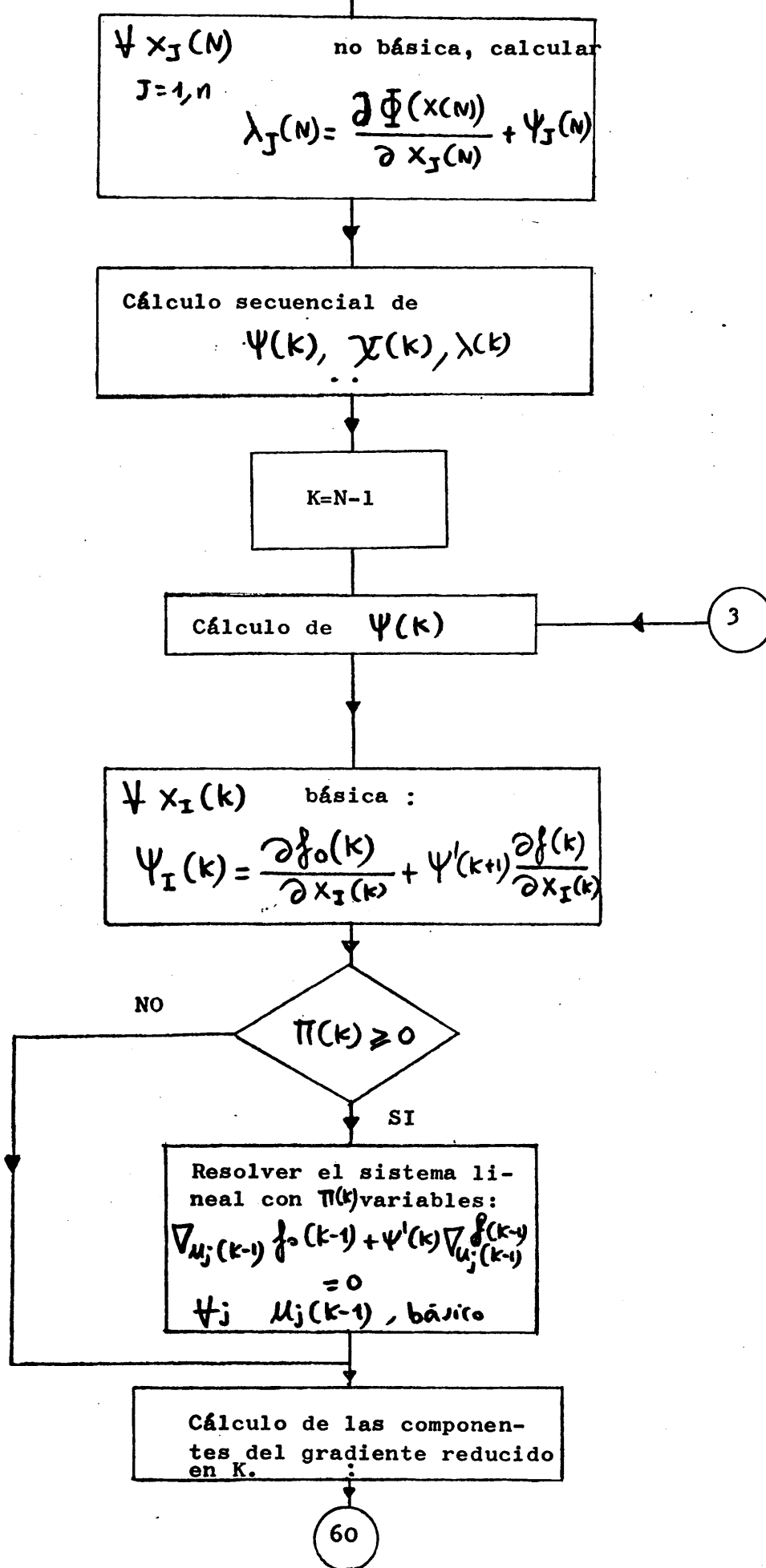
$$i \in VNB(k), j \in CB(k)$$

o es posi-
le efectuar
1 cálculo
ecuencial.

1







$$\forall \mu_I(k-1) \quad \text{no básica:}$$

$$y_I(k-1) = \frac{\partial f_0(k-1)}{\partial \mu_I(k-1)} + \psi'(k) \frac{\partial f(k-1)}{\partial \mu_I(k-1)}$$

$$\forall x_I(k) \quad \text{no básica:}$$

$$\lambda_I(k) = \frac{\partial f_0(k)}{\partial x_I(k)} + \psi'(k+1) \frac{\partial f(k)}{\partial x_I(k)} - \psi_I(k)$$

K=K-1

NO

K=0

SI

FIN del cálculo
del gradiente
reducido.

Test de Optimalidad
sobre las componentes
del gradiente reducido
proyectado

FIN

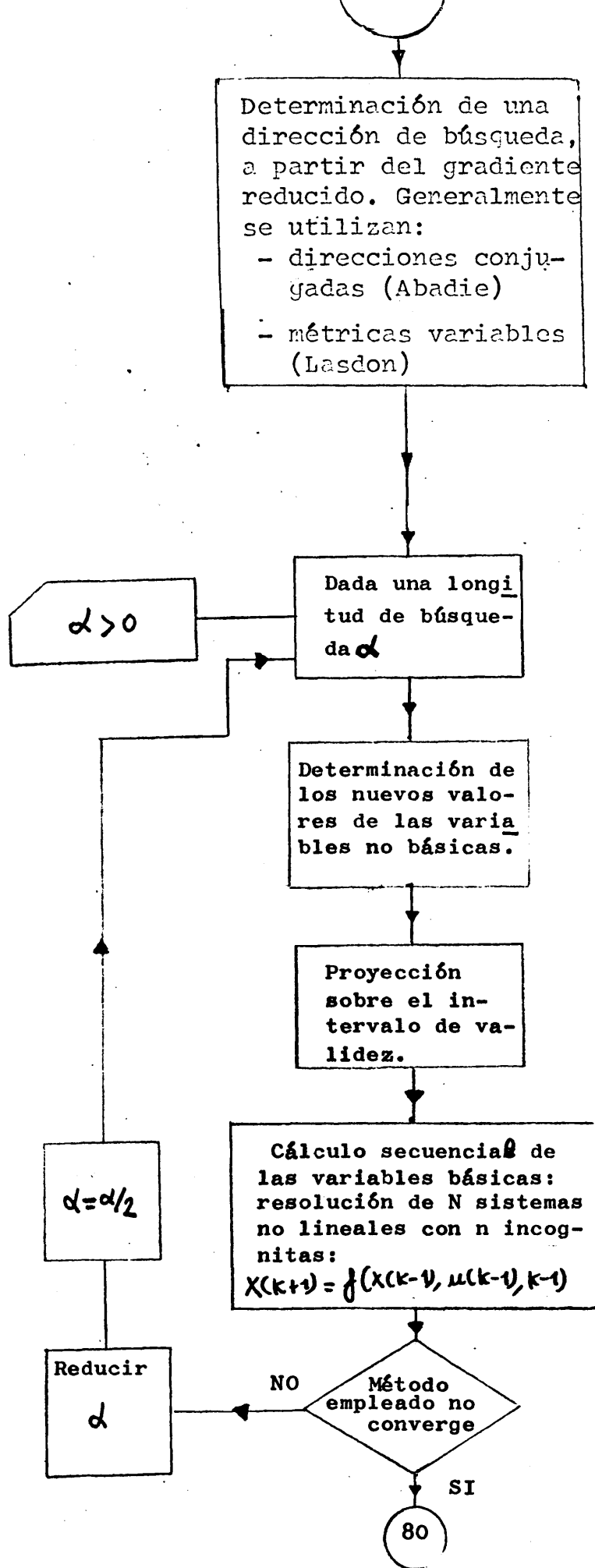
La itera-
ción actual
contiene
la solu-
ción.

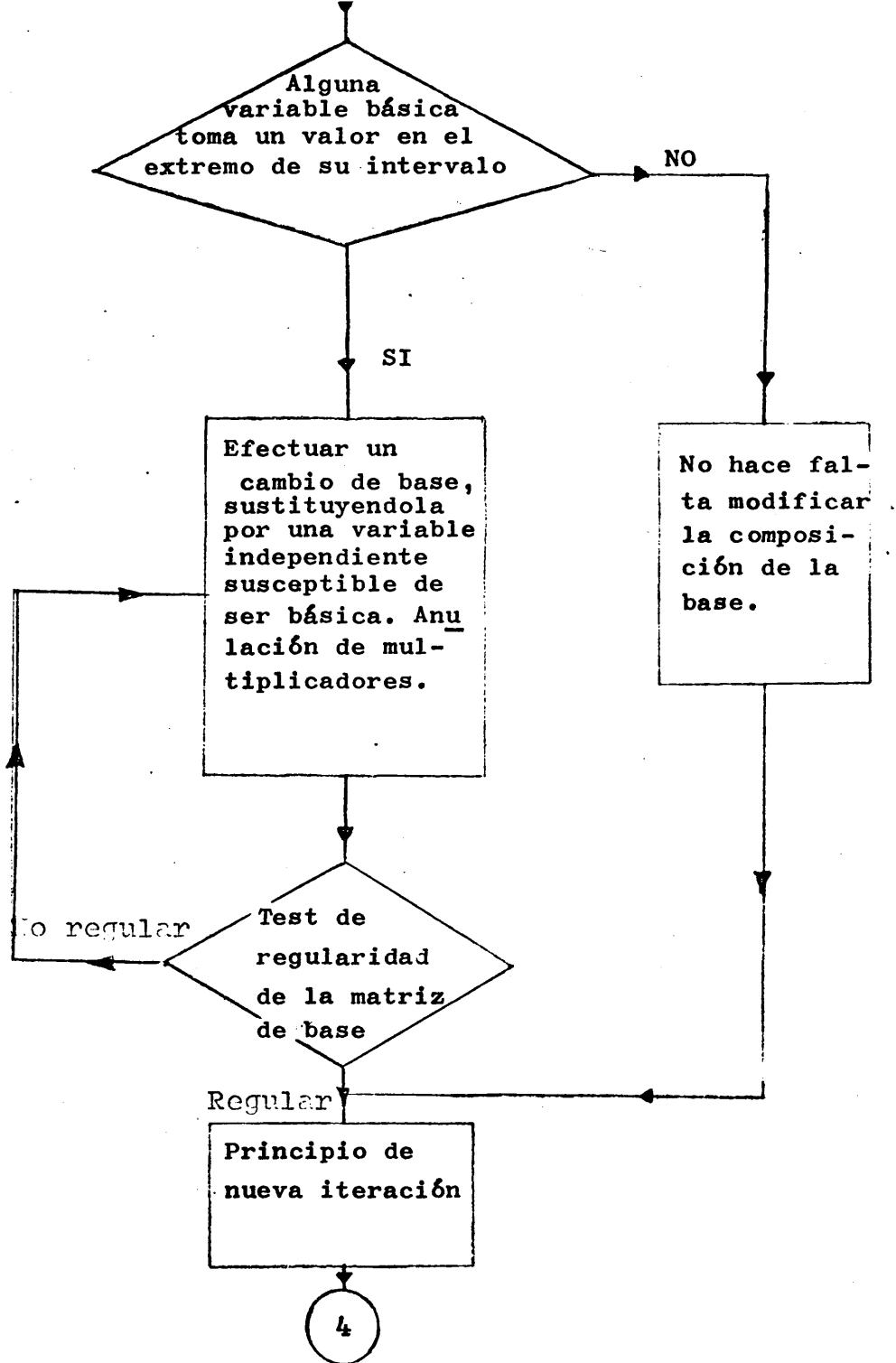
SI

Gradien-
te reducido
proyectado
nulo

NO

70





Conociendo una solución factible inicial, el algoritmo descrito calcula eficientemente el gradiente reducido, aprovechando la estructura del jacobiano de las restricciones de un problema de control y genera una serie de soluciones factibles monotonamente mejoradas. En el cálculo del gradiente reducido se obtienen simultáneamente los valores de las variables duales.

Sin embargo, como lo ilustra el anterior organigrama, el proceso de cálculo del gradiente reducido es numéricamente complejo, a lo que hay que añadir la dificultad que presente la solución del problema reducido.

Por otra parte, las únicas restricciones instantáneas sobre la trayectoria que el método descrito considera, son las de cotas en las variaciones individuales de las variables de control y/o estado. Reconociendo los importantes condicionamientos reales que tales restricciones traducen, no es menos evidente que la consideración de restricciones instantáneas conjuntas control/estado que hagan intervenir simultáneamente más de una variable, resultan imprescindibles para modelar situaciones reales, como veremos en el capítulo IV.

Para aplicar la metodología del GRG, tales restricciones deben convertirse en igualdades mediante la introducción de variables de holgura con la restricción individual asociada de positividad:

$$g(x(k), u(k)) \leq 0 \longrightarrow g(x(k), u(k)) + H(k) = 0$$

$$H(k) \geq 0$$

Sin embargo, existe una diferencia básica entre las restricciones representadas por las ecuaciones de estado y

las del tipo $g(x(k), u(k)) \leq 0$

.Mientras que

las primeras son, por definición, constantemente activas, las segundas pueden ser activas o no, según cual sea la trayectoria que se esté considerando y por lo tanto algunas de ellas podrían no ser tomadas en consideración en las distintas iteraciones.

Esta consideración esta incorporada en el código GRG elaborado en la Case Western Reserve University por el grupo del profesor LASDON (117, 118) Este, junto con el código GREG de Electricidad de Francia elaborado por el grupo del profesor Abadie (115, 116) constituyen los más significativos instrumentos prácticos que aplican los métodos GRG.

Al menos en las versiones que nos son conocidas, ambos están diseñados para su aplicación a programas matemáticos de estructura general y no utilizan el cálculo secuencial del gradiente reducido aprovechando la estructura de un problema de control. Sin embargo, Abadie y Robert han utilizado un código que instrumenta este procedimiento, por él propuesto, en la resolución del problema de Kendrick y Taylor como expondre mos en el capítulo IV.

Las diferencias entre ambos códigos se sitúan en la forma en que resuelven los problemas reducidos y en la consideración exclusiva o no de las únicas restricciones activas.

La exposición detallada de los procedimientos utilizados desborda los temas de nuestro interés y está perfectamente expuesto en las referencias (117, 118) Basta señalar que :

- a) el método de Lasdon utiliza un algoritmo de métrica variable y el de Abadie el del gradiente o gradiente conjugado opcionalmente.

b) en el código GREG, se efectúa un cambio de base cada vez que una variable básica viola uno de sus límites en el transcurso de las iteraciones por el método de Newton para regresar a la superficie de las restricciones. Ello puede dar lugar a cambios de base tales que produzcan un empeoramiento en el valor de la función objetivo con respecto al de la anterior iteración. Esta dificultad, señalada por Lasdon, esta subsanada en el código de la Case Western Reserve University.

En resumen pues, GRG es un método primal que genera información dual eficientemente. Se inscribe así en la primera de las dos líneas de evolución apuntadas en la introducción a este capítulo.

El elevado coste de mantener la factibilidad de todas las soluciones, se justifica cuando las restricciones del problema y sus ecuaciones de estado se conocen con un grado tal de seguridad que carece de interés la solución de problemas que difieren poco del planteado.

Este no es precisamente el caso de las aplicaciones económicas. En estas, la imprecisión en los datos hace que, más que un problema específico, nos interese toda una familia de problemas con análoga estructura. En otras palabras, más que la solución en sí misma, nos interesan indicadores de su sensibilidad frente a variaciones en los datos del problema.

Con este planteamiento, los métodos duales cobran especial importancia. Por la relativa sencillez de su proceso de cálculo y lo ilustrativo que este resulta desde un punto

de vista económico, porque abren el camino al estudio de los procesos de descentralización-descomposición, por los indicadores de sensibilidad que generan y porque no precisan el conocimiento de una solución inicial factible, los métodos duales y los intentos de extender su campo de aplicación, aparecen como especialmente interesantes. Este es el objeto del siguiente apartado y del capítulo V.

III.4.1- INTRODUCCIÓN A LOS MÉTODOS DUALES.

En esta introducción resumimos brevemente la relación entre convexidad total o local y separabilidad en los métodos duales. Al mismo tiempo encuadramos entre ellos al método de los multiplicadores de Héstenes.

Entendemos que la aplicación de los métodos duales al programa matemático:

$$\text{Min } J(x) \quad \left\{ x \equiv (x(k), u(k-1), k=1, \dots, N) \right\}$$

$g(x) = 0$ (restricciones igualdad representadas por las ecuaciones de estado).

$r(x) \leq 0$ (restricciones sobre la trayectoria).

formulado a partir de un problema discreto de control óptimo está dominado por dos conceptos: Convexidad y separabilidad. El primero es una propiedad que puede o no poseer el problema y el segundo es su consecuencia en la operativa de resolución.

La solución del problema dual:

$$\text{Max}_{\lambda, \mu \geq 0} \Phi(\lambda, \mu)$$

donde $\Phi(\lambda, \mu)$ es la función dual

$$\Phi(\lambda, \mu) = \min_x [J(x) + \mu' r(x) + \lambda' g(x)]$$

o bien

$$\max_{\lambda} \Phi(\lambda)$$

con

$$\Phi(\lambda) = \min_{\substack{x \\ r(x) \leq 0}} [J(x) + \lambda' g(x)]$$

~~si utilizamos dualidad parcial, tomará el mismo valor que~~
el problema primal:

$$\max_{\lambda, \mu \geq 0} \Phi(\lambda, \mu) = \min_{\substack{x \\ g(x)=0, r(x) \leq 0}} J(x)$$

si el Lagrangiano posee un punto de silla. La existencia del mismo está garantizada si el programa matemático formulado es un programa convexo. De no ser así nada garantiza la equivalencia de ambos problemas.

Recuerdese sin embargo que, con independencia de la convexidad del problema primal, la función dual es cóncava y representa una cota inferior a los valores tomados por la función ~~objetiva~~ del primal. Lardon (L5)

En programas no convexos, los métodos duales pueden aplicarse si el punto solución verifica la hipótesis de convexidad local: en dicho punto el Hessiano del Lagrangiano no es definido positivo en todo el espacio (y no solamente en el subespacio tangente a las restricciones como la exige su condición de solución local).

Si ni siquiera la hipótesis de convexidad local resulta cierta, las soluciones de los problemas primal y dual no coincidirán en general y los métodos duales son inaplicables. Las referencias (L5) (Lasdon), (W1) (Wisner), (L13) (Luenberger) contienen una completa exposición de este tema.

Puesto que la minimización del Lagrangiano como una función explícita de las variables duales es poco eficiente numericamente, la búsqueda de la solución se plantea como un proceso secuencial a dos niveles. El primer nivel calcula la función dual para unos valores de los multiplicadores dados por el segundo nivel. Este los calcula a través de la maximización de la función dual con los valores de x resultantes del proceso en el primer nivel. La importancia del método para problemas de control como los que nos ocupan es doble:

- a) porque conduce a una separación del problema primal en una serie de problemas independientes. En ello se basan eficientes algoritmos tipo Pearson (P4) y Tamura (T3) y Lasdon (L5).
- b) porque permite abordar problemas cuya dinámica presente efectos retardados en el tiempo sin necesidad de aumentar el espacio de estado.

Conceptualmente, la interpretación del papel jugado por las variables en el proceso se inscribe en la línea descrita en el capítulo II y será explotada en los capítulos siguientes.

La convexidad del problema, o en su defecto la hipótesis de convexidad local, permite desarrollar métodos duales

que se revelan eficientes en problemas cuya formulación incluya las fuentes de mayor complejidad computacional : restricciones sobre la trayectoria y retrasos en el tiempo en las ecuaciones del estado.

Si tal hipótesis no se cumple, la formulación de Héstenes (H5) de problemas con restricciones igualdad :

$$\begin{aligned} \text{Min } J(x) \\ g(x)=0 \end{aligned}$$

hace que el Hessiano del Lagangiano del problema modificado:

$$\begin{aligned} \text{Min } [J(x) + K g'(x) g(x)] \\ g(x)=0, K > 0 \end{aligned}$$

sea, en los óptimos locales del problema, definido positivo en todo el espacio. En consecuencia es posible aplicar al problema modificado la teoría de la dualidad local.

Aunque el método de los multiplicadores de Hestenes (o "augmented penalty function") fué planteado por su autor sin ninguna relación con los métodos duales, esta ha sido así evidenciada por Luenberger(L13).

El objetivo perseguido por Héstenes era simplemente salvar la imprecisión con la que se obtienen, en el óptimo, las variables duales cuando se aplica un método de penalización. Como se sabe, utilizando tal método se obtienen los multiplicadores como el producto de un factor que tiende a cero por uno que tiende a infinito.

El método de Héstenes es también un proceso iterativo a dos niveles, con una expresión para actualizar los multiplicadores:

$$\lambda(k+1) = \lambda(k) + 2 K g(x)$$

que es la misma que se obtiene cuando se aplica el método de Newton a la maximización de la función dual teniendo en cuenta la forma $\frac{1}{2K} I$ a la que tiende su Hessiano al aumentar la constante de penalización K .

El método, no solamente permite la aplicación de la dualidad parcial, sino que además la favorece. En efecto, el proceso de maximización de la función dual se ve facilitado por la acción conjunta de tres circunstancias:

- a) Su concavidad
- b) La simple expresión de su gradiente : $g(x)$
- c) La favorable estructura de autovalores de su Hessiano.

Sin embargo, la formulación de Héstenes destruye la separabilidad en la resolución que podría obtenerse aplicando, si ello fuera posible, los métodos duales puros al programa matemático obtenido de un problema discreto de control óptimo.

Demostraremos en su momento que ello no quita viabilidad práctica al método en el caso en que no existan excesivas restricciones instantáneas sobre la trayectoria y se introduzcan las modificaciones apuntadas por Miele (M8) para acelerar su convergencia. La existencia de efectos retardados en las ecuaciones de estado no debe de afectar excesivamente sus requerimientos en memoria y tiempo de cálculo.

En resumen:

- 1) En problemas que satisfagan la hipótesis de convexidad local, los métodos duales permiten la descomposición de la fase primal en el proceso iterativo de cálculo. Gracias a ello son eficientes en los casos que incluyen restricciones sobre la trayectoria y retrasos en las ecuaciones de estado. El apartado siguiente expone uno de estos métodos del tipo Lasdon-Tanura en el que utilizamos dualidad parcial y el método del gradiente reducido en la solución de los problemas locales.
- 2) Aunque la hipótesis de convexidad local no se satisfaga en el problema original, si que se verifica en el problema modificado de Héstenes. Este puede ser concebido como un método dual a dos niveles. El método produce una pérdida de separabilidad en la solución de un problema de control. A pesar de ello, su eficacia puede mantenerse, aún en la presencia de efectos retardados, si se prescinde de excesivas restricciones sobre la trayectoria. A esta cuestión dedicamos el punto III. 4. 2..

III. 4. 2. ALGORITMO BASADO EN EL METODO DUAL DE LOS
MULTIPLICADORES DE HESTENES.

En su versión original, el método propuesto por Hestenes substituye el problema :

$$\begin{aligned} \min J(x) \\ g(x) &= 0 \\ x &\in E^n \\ E^n &\xrightarrow{g} E^q, \quad q < n \end{aligned}$$

por el equivalente :

$$\begin{aligned} \min J(x) + K g'(x) g(x) \\ g(x) &= 0 \\ K &> 0 \end{aligned}$$

Referencias detalladas sobre el método de Hestenes pueden encontrarse en Hestenes (H5), Miele (M8), y Rockafellar (R6). En su aplicación práctica, lo resume el organigrama de la figura FIII.4

Mas que la justificación original dada por su autor, directamente relacionada con los métodos de penalización, nos interesa analizar su relación con los métodos duales.

Observe que las condiciones de primer orden satisfechas en un punto solución (x^*, λ^*) son las mismas para el problema original y el modificado:

$$\nabla_x J(x^*) + \lambda^{*'} \nabla_x g(x^*) = 0$$

En (x^*, λ^*) , el Hessiano del Lagrangiano del problema original debe ser semidefinido positivo en el subespacio tangente a las restricciones, T :

$$T \equiv \{x \mid \nabla g(x^*)x = 0\}$$

Si la condición suficiente de 2^o orden se verifica, debe de ser definido positivo en dicho subespacio. Pero nada le obliga a serlo en todo el espacio, y de no serlo, la teoría de la dualidad local no podría aplicarse.

Sin embargo, el Hessiano del Lagrangiano del problema modificado se obtiene sumando al del problema original, $H \mathcal{L}(x^*, \lambda^*)$, la expresión $2K[\nabla g(x^*)]' \nabla g(x^*)$:

$$H \mathcal{L} M(x^*, \lambda^*) = H \mathcal{L}(x^*, \lambda^*) + 2K[\nabla g(x^*)]' \nabla g(x^*)$$

La hipótesis de convexidad local se satisface ahora para adecuados valores de K, puesto que al ser $H \mathcal{L}(x^*, \lambda^*)$ definido positivo en T y $[\nabla g(x^*)]' \nabla g(x^*)$ en su complementario, su suma lo será en todo el espacio.

La aplicación del método dual al problema modificado lleva consigo:

- 1) Evaluación de la función dual para λ, K dados:

$$\Phi(\lambda) = \min_x [J(x) + \lambda' g(x) + K g'(x) g(x)]$$

Este paso se corresponde con el módulo 1 del organigrama del método de Héstenes (figura EIII.4).

- 2) Maximización de la función dual

$$\max_{\lambda} \Phi(\lambda)$$

con x tomando los valores resultantes del proceso anterior.

Este paso puede realizarse de distintas formas, de cada una de las cuales resultará una expresión secuencial de modificación de los multiplicadores. Si nos propusiéramos emplear el método de Newton, podríamos aprovechar dos resultados :

- que el Hessiano de la función dual tiende a $\frac{1}{2K} I$ para valores crecientes de K. (Luenberger, (AB))
- que el gradiente de la función dual, (en las con-

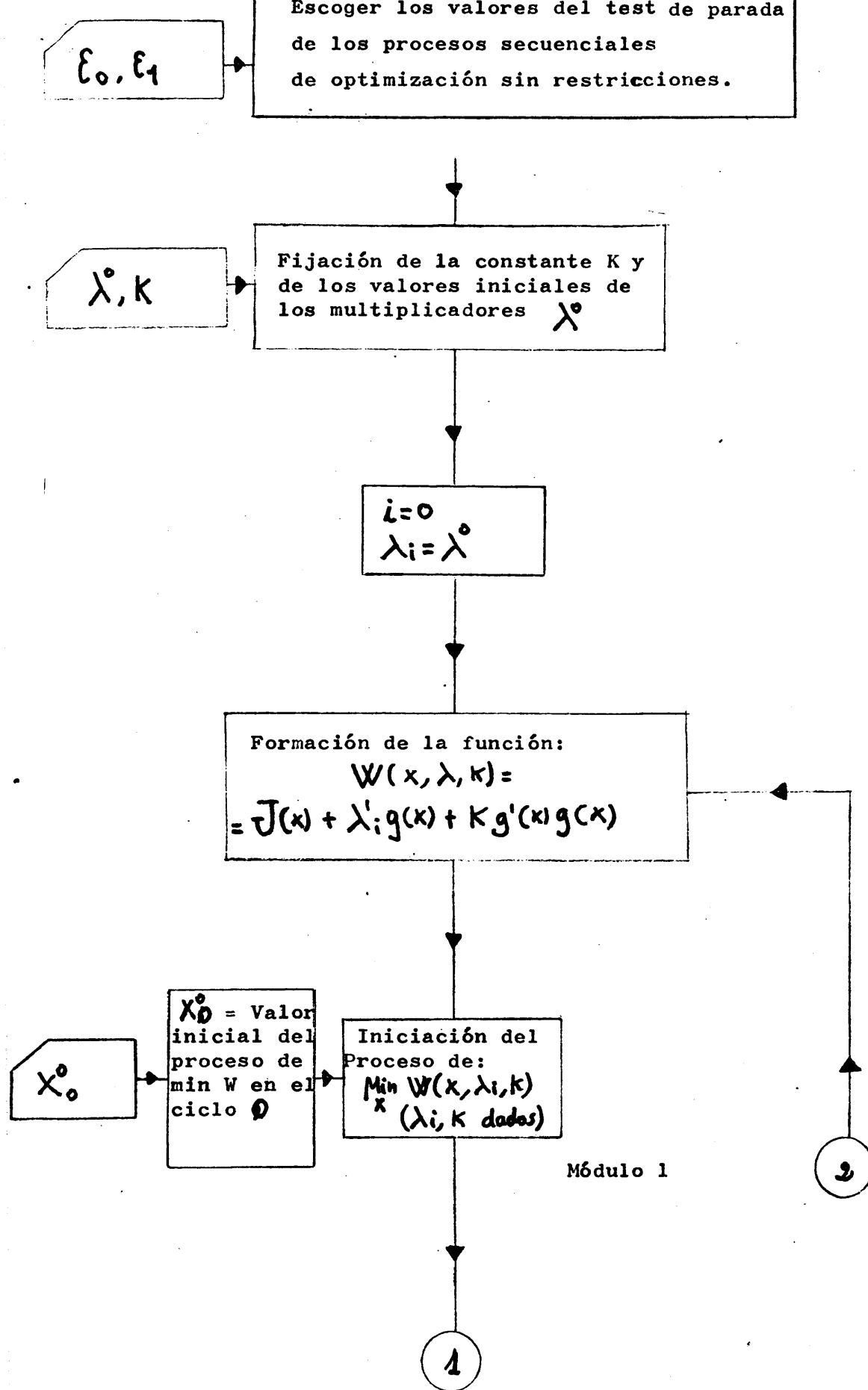


Figura FIII.4: Organigrama de realización del método de Hestenes

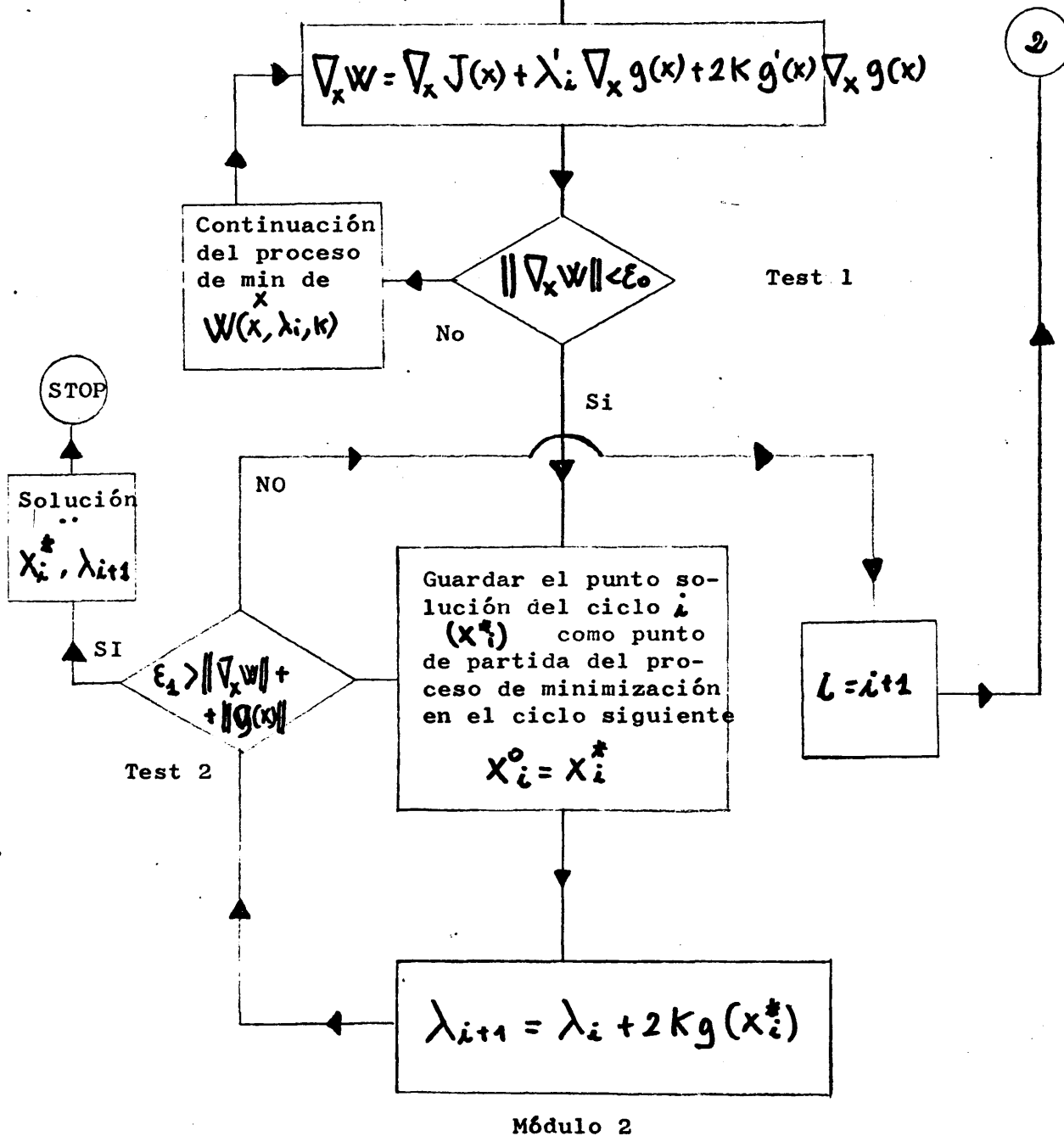


Figura FIII.4: Organigrama de realización del método de Hestenes
(continuación)

diciones dadas en la referencia), viene dado por el error cometido en la satisfacción de las restricciones igualdad. (Lasdon (L5)).

$$\nabla_{\lambda} \Phi(\lambda) = g'(x)$$

En tal caso, la expresión secuencial del método de Newton (Apéndice III) vendría dada por :

$$\lambda_{i+1} = \lambda_i + 2k g(x)$$

Dicha expresión es la utilizada por Héstenes para la actualización de los multiplicadores (modelo 2, fig. FIII.4). Este es pues un método dual utilizando un método de Newton modificado (Hessiano constante), en la maximización de la función dual.

Observese que:

- 1) la constante K permanece constante durante todo el proceso.
- 2) los multiplicadores sólo son modificados al final de cada ciclo. En cada uno de ellos se efectúa un proceso completo de minimización de $\Psi(x, \lambda, k)$.

El reconocimiento de su carácter de método dual aplicado a un problema localmente convexo sugiere modificaciones al método de Hestenes especialmente interesantes en su aplicación a problemas discretos de control óptimo con efectos retardados en las ecuaciones de estado, que exponemos a continuación :

1. Substitución del método de Newton por el del gradiente en el cálculo de $\underset{\lambda}{\text{Max}} \Phi(\lambda)$

Puesto que la función dual es cóncava en cualquier subconjunto convexo de su dominio de definición, y

dada la favorable estructura de autovalores de su Hessiano para adecuados valores de K , las condiciones para una rápida y adecuada convergencia del método del gradiente parecen reunidas.

Por otra parte, el método de Héstenes no es sino una aproximación al método de Newton.

La modificación de los multiplicadores se hará de acuerdo con la expresión :

$$\lambda_{i+1} = \lambda_i + \alpha \nabla_{\lambda}' \Phi(\lambda) = \lambda_i + \alpha g(x)$$

El valor de α es el correspondiente al punto máximo, en la dirección de búsqueda así fijada, de la función tomada como función de descenso. Normalmente esta es la misma función objetivo $\Phi(\lambda)$. Pero la dependencia implícita de x^* con respecto a λ hace difícil la determinación de α . En cambio, si escogemos como función de descenso la condición de optimalidad en el problema original (Miele (M8)) :

$$F_x(x, \lambda) = \nabla_x J(x) + \lambda' \nabla_x g(x)$$

$$F_x(x^*, \lambda^*) = 0$$

es posible llegar a una expresión para α :

$$\alpha = \arg \min F_x(x, \lambda_{i+1}) F_x'(x, \lambda_{i+1}) = \arg \min \| F_x(x, \lambda_{i+1}) \|^2$$

$$F_x(x, \lambda_{i+1}) = \nabla_x J(x) + \lambda_{i+1}' \nabla_x g(x) = \nabla_x J(x) + (\lambda_i + \alpha g(x))' \nabla_x g(x)$$

$$= \nabla_x J(x) + \lambda_i' \nabla_x g(x) + \alpha g'(x) \nabla_x g(x)$$

$$= F_x(x, \lambda_i) + \alpha g'(x) \nabla_x g(x)$$

$$F_x(x, \lambda_{i+1}) F_x'(x, \lambda_{i+1}) =$$

$$= [F_x(x, \lambda_i) + \alpha g'(x) \nabla_x g(x)] [F_x'(x, \lambda_i) + \alpha \nabla_x' g(x) g(x)] =$$

$$\begin{aligned}
&= F_x(x, \lambda_i) F'_x(x, \lambda_i) + \alpha F_x(x, \lambda_i) [\nabla_x g(x)]' g(x) + \\
&\quad + \alpha g'(x) \nabla_x g(x) F'_x(x, \lambda_i) + \\
&\quad + \alpha^2 g'(x) \nabla_x g(x) [\nabla_x g(x)]' g(x)
\end{aligned}$$

$$\begin{aligned}
\frac{\partial \|F_x(x, \lambda_{i+1})\|}{\partial \alpha} &= F_x(x, \lambda_i) [\nabla g(x)]' g(x) + \\
&\quad + g'(x) \nabla g(x) F'_x(x, \lambda_i) + \\
&\quad + 2\alpha g'(x) \nabla_x g(x) [\nabla g(x)]' g(x)
\end{aligned}$$

En donde :

llamando: $[\nabla_x g(x)]' g(x) = P_x$

$$\frac{\partial \|F_x(x, \lambda_{i+1})\|}{\partial \alpha} = 2 F_x P_x + 2\alpha P'_x P_x = 0$$

Se obtiene

$$\alpha = - \frac{F_x P_x}{P'_x P_x}$$

Valor que minimiza realmente
puesto que

$$\|F_x(x, \lambda_{i+1})\|$$

$$\frac{\partial^2 \|F_x(x, \lambda_{i+1})\|}{\partial \alpha^2} = 2 P'_x P_x > 0, \forall P_x \neq 0$$

2. Modificación de los multiplicadores, de acuerdo con la fórmula expresada, a cada iteración en la evaluación de la función dual.

No encontramos una razón teórica que garantice una aceleración en la convergencia. Pero parece intuitivamente razonable alternar los pasos iterativos del método del

- gradiente aplicado a la evaluación de la función dual y su maximización. Al así hacerlo se aprovecha inmediatamente toda mejora obtenida en los valores de X para el cálculo de λ y viceversa. Ello equivaldría a suponer reducida a una única iteración la longitud del ciclo en el método de Héstenes y suprimir el test 1 en su proceso.
3. Modificar secuencialmente la constante de penalización K. Esta es mantenida constante en el método de Hestenes. Pero puesto que las variaciones en x en cada paso iterativo producen variaciones en el grado de satisfacción de las restricciones ($g(x)=0$) que se pueden expresar como una función de K , esta puede ser escogida de forma que, después de cada iteración, las restricciones se satisfagan (en media) hasta el primer orden. Un razonamiento mas detallado y técnico puede encontrarse en la referencia Miele (124). Aceptando su resultado, el valor de K lo fijamos a través de la expresión :

$$K = \frac{g'(x) g(x)}{2 P'_x P_x}$$

en la cual K permanece finita en la convergencia.

El algoritmo resultante de estas modificaciones es el que describe el organigrama de la figura 5, al que llamamos método dual de los multiplicadores.

El proceso de cálculo así resumido se limita básicamente a un conjunto de multiplicaciones matriciales , sin necesidad de efectuar ninguna inversión. Se observará su relativa sencillez comparativamente al método primal basado en el gradiente reducido generalizado. Ello es debido fundamentalmente a la no consideración de restricciones instantáneas sobre la trayectoria.

culo.

Se entiende por solución al punto x^*, λ^* que verifique las restricciones y la condición de optimalidad de primer orden, con las tolerancias ϵ_1, ϵ_2 :

$$\|\nabla_x J(x^*) + \lambda^{*'} \nabla_x g(x^*)\| < \epsilon_1$$
$$\|g(x)\| < \epsilon_2$$

Propuesta dual inicial (valores iniciales de los multiplicadores)

$\lambda = \lambda^0$

Valor inicial (arbitrario) de la constante de penalización

$K = K^0$

Punto origen de las iteraciones por el método del gradiente en la evaluación de la función dual

x^0
 $g(x^0) \neq 0$

10

Cálculo de la primera propuesta primal :

Calcular: $g(x^0)$
 $\nabla_x g(x^0)$
 $\nabla_x J(x^0)$
 $P_x = \nabla_x' g(x^0) g(x^0)$
 $F_x = \nabla_x J(x^0) + \lambda' \nabla_x g(x^0)$
 $\nabla_x W(x^0, \lambda) = F_x(x^0, \lambda) + 2K P_x'$

1

$x = x^0 - \beta \nabla_x W(x^0, \lambda)$
 $\beta = \arg \min W(x, \lambda)$

Vector primal-dual
candidato a solución

x, λ

Cálculo de los indi-
cadores de test de
parada-

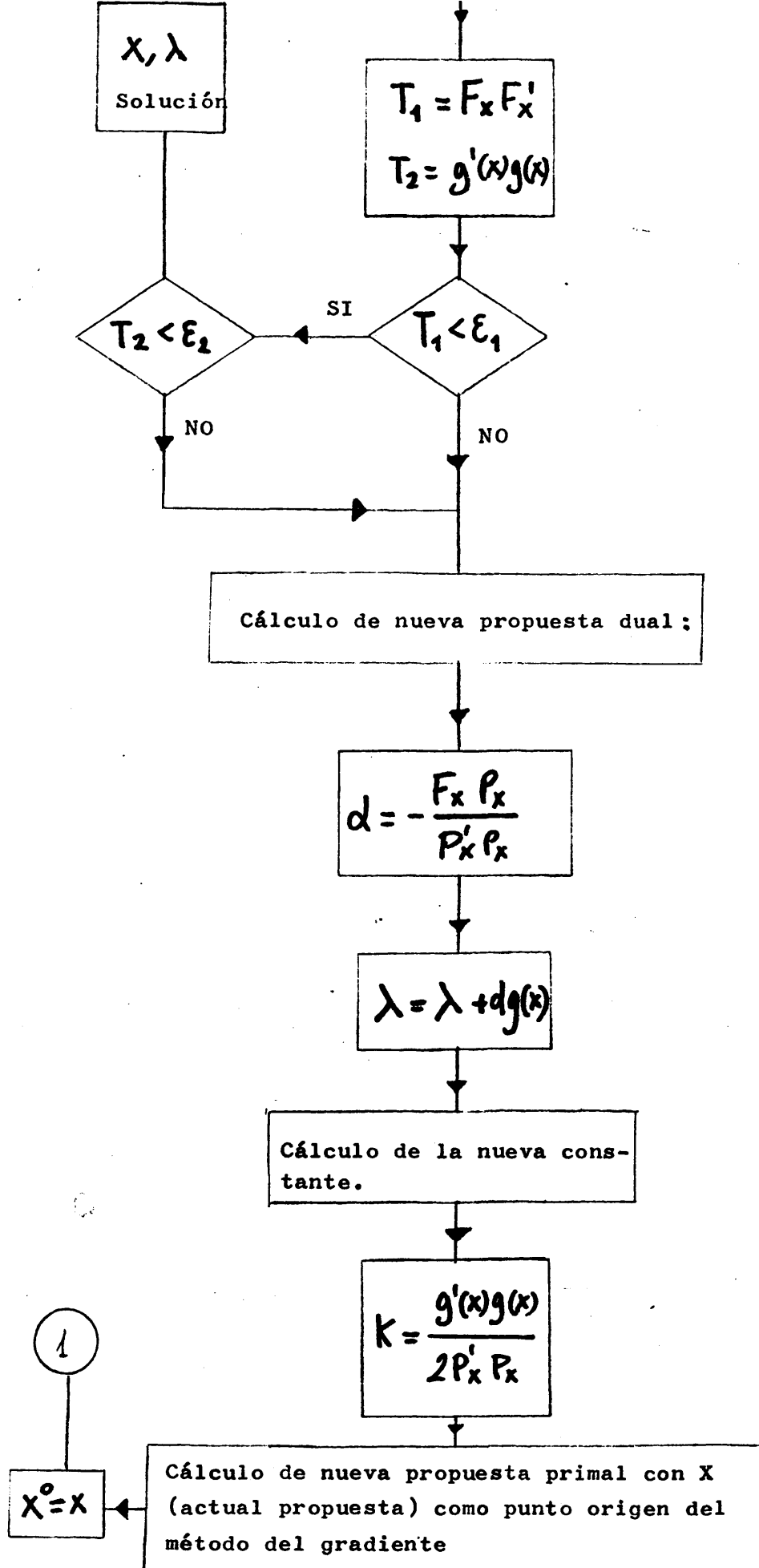
Calcular:

$g(x), \nabla_x g(x), \nabla_x J(x)$

$P_x = \nabla_x' g(x) g(x)$

$F_x = \nabla_x J(x) + \lambda' \nabla_x g(x)$

20



• Recuerdese que dicho método primal se basa en una estructura particular del jacobiano de las restricciones que desaparece si la dinámica del sistema analizado presenta retrasos en el tiempo. En cambio, la presencia de los mismos no afecta para nada la viabilidad del método dual de los multiplicadores puesto que no está basado en absoluto en la estructura de la matriz

Ello no quiere decir que no la tengamos en cuenta en el momento de programar la ejecución del cálculo. Lo importante es que, cualesquiera que sean las modificaciones que introduzcan los efectos retardados, no imposibilitan la aplicación del método.

Analícemos tal aplicación cuando el problema al que se aplica el proceso es uno de control óptimo discreto. Desarrollamos a continuación dos series de consideraciones relativas a:

- la capacidad de memoria rápida requerida por el método y los condicionamientos que ello impone en el desarrollo del cálculo.
- la estructura de la matriz $\nabla_x g(x)$ y la aceleración del proceso de cálculo resultante cuando esta se toma en cuenta.

Con respecto al primer punto, analicemos las dimensiones de los elementos del organigrama en el caso de un problema de control óptimo que se extienda a N periodos de tiempo ($K = 1, N$) con n variables de estado y m de control (Para fijar ideas supongamos : a) $N=30, n=6, m=8$

b) $N=30, n=10, m=15$

El número total de variables a considerar sería de $Nn + (N-1)m$, o suponiendo que el estado inicial está fijado y deja de considerarse como una variable $(N-1)(n+m)$

[a) 406, b) 725]

El número total de restricciones igualdad : $N.n$

[a) 180, b) 300.]
Las dimensiones de los distintos elementos utilizados son:

		a)	b)
$g(x)$	$(N.n \times 1)$	(180×1)	(300×1)
$\nabla_x J(x)$	$(1 \times (N-1)(n+m))$	(1×406)	(1×725)
x	$((N-1)(n+m) \times 1)$	(1×406)	(1×725)
λ	$(N \times 1)$	(180×1)	(300×1)
F_x	$(1 \times (N-1)(n+m))$	(1×406)	(1×725)
$\nabla_x w$	$(1 \times (N-1)(n+m))$	(406×1)	(725×1)
P_x	$((N-1)(n+m) \times 1)$	(406×1)	(725×1)

Es decir un total de $5(N-1)(n+m) + 2N.n$, [a) 2360, b) 3225]

Es perfectamente posible disponer de tal cantidad de registros en memoria rápida. La dificultad proviene del elemento todavía no contabilizado $\nabla_x g(x)$. Sus dimensiones son $(N.n \times (N-1)(n+m))$, [a) $(780 \times 406) = 73080$, b) $(300 \times 725) = 217.500$],., cantidades absolutamente fuera de los límites disponibles.

El problema se soluciona inmediatamente gracias a dos observaciones, una relativa a la organización del cálculo, la otra a la estructura de dicho elemento.

Cada una bastaría por sí sola a salvar la dificultad. Su acción conjunta reduce la memoria precisada por las operaciones con $\nabla_x g(x)$ a una cantidad menor de $N.n$, [180, 300], al tiempo que debe disminuir notablemente el tiempo de cálculo. En efecto: Por una parte, nada obliga a calcular en bloque y mantener en memoria toda la matriz $\nabla_x g(x)$. Las únicas operaciones a efectuar con ella son $[\nabla_x g(x)]' g(x)$ en el cálculo de P_x y $\lambda' \nabla_x g(x)$ en el cálculo de F_x .

Tanto F_x como P_x pueden ser calculados componente a componente, para lo cual no se precisa mas que la disponibilidad de las sucesivas correspondientes columnas de $\nabla_x g(x)$, es decir $N.n, [(180, 300)]$ posiciones suplementarias de memoria.

Por otra parte, la estructura dinámica de las restricciones da a la matriz $\nabla_x g(x)$ la particularísima forma expuesta en III. 3. Ello implica que somos capaces de delimitar a priori amplias zonas que sabemos constituidas por ceros. La parte de las sucesivas columnas consideradas que pertenezcan a dichas zonas no deberán ser calculadas, almacenadas en memoria ni utilizadas en el cálculo.

Suponiendo que no hubiera efectos retardados en las ecuaciones de estado, la forma de $\nabla_x g(x)$ sería:

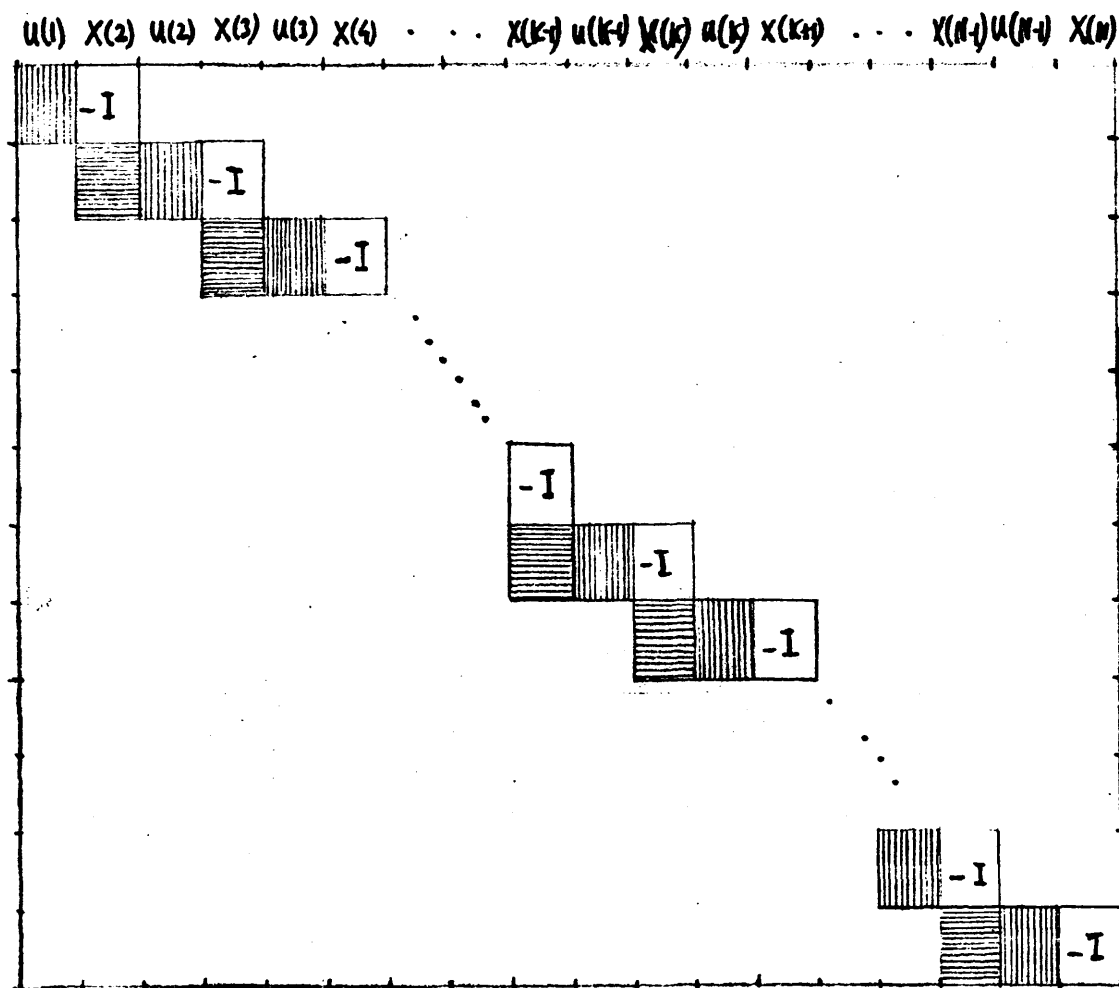


FIGURA FIII.6 ESTRUCTURA DEL JACOBIANO DE LAS ECUACIONES DE ESTADO CUANDO NO HAY EFECTOS RETARDADOS.

Es decir, cada columna tendría como máximo $n + 1$ elementos no nulos.

Suponiendo la existencia de retrasos en las ecuaciones de estado :

$$x(k+1) = f(x(k), x(k-1), \dots, x(k-l), u(k), u(k-1), \dots, u(k-p))$$

Sea $\Theta = \max [l, p]$ el mayor efecto retardado contenido en las ecuaciones. La forma de la matriz sería entonces, suponiendo $l=3, p=2$

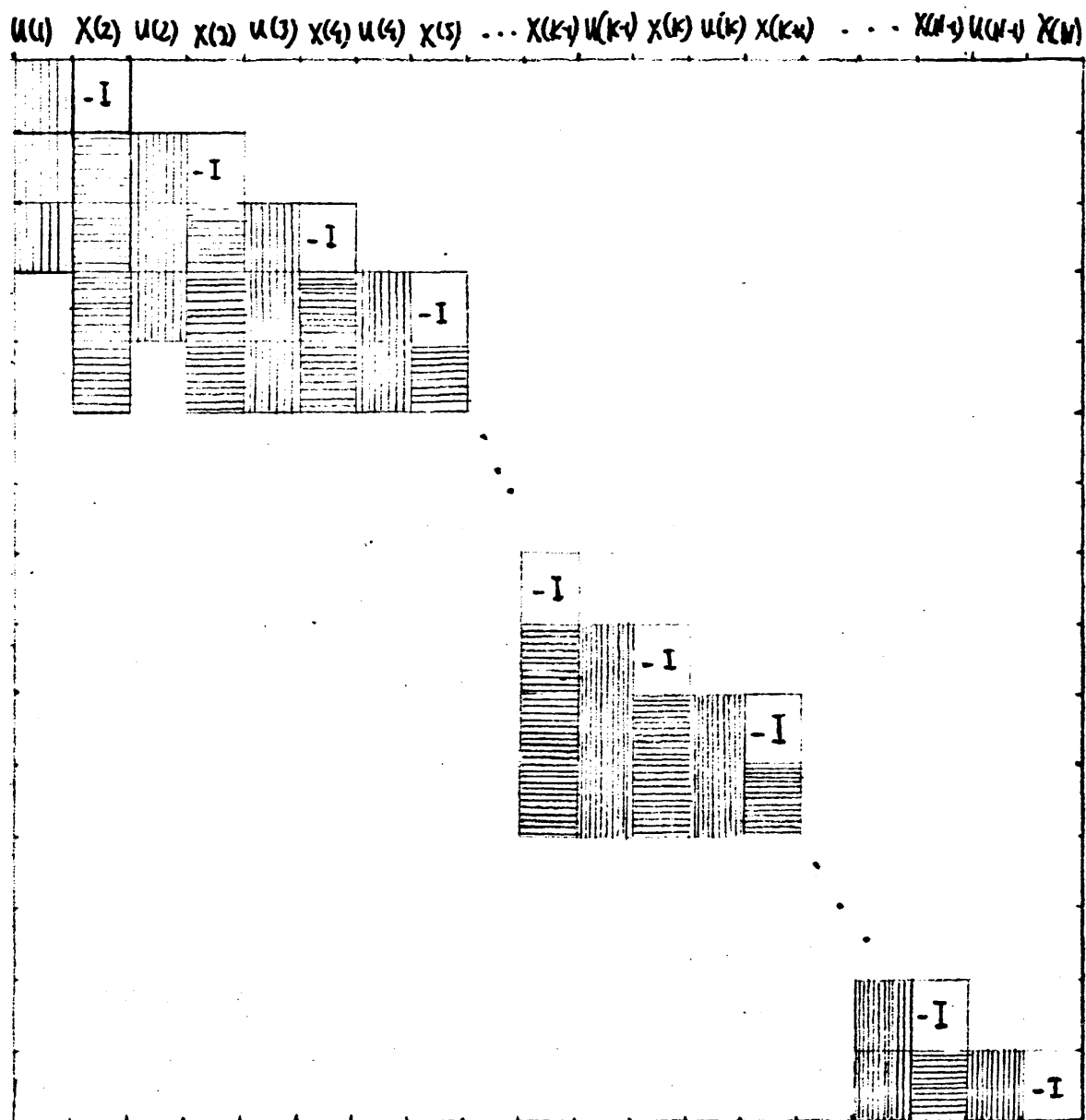


FIGURA FIII.7 ESTRUCTURA DEL JACOBIANO DE LAS ECUACIONES DE ESTADO CUANDO EXISTEN EFECTOS RETARDADOS EN ESTADOS Y CONTROLES.

Luego el máximo número de elementos no nulos a priori en cada columna es $(0+1)n+1$. En cualquier caso pues, el número de registros de memoria necesarios simultaneamente para operaciones con $\nabla_x g(x)$ es menor de $N.n$.

Si es posible instrumentar un modelo que no precise explícitas restricciones sobre los valores instantáneos de las variables de estado y/o control, este método parece ser eficaz y facil de aplicar. La limitación que queda por analizar es la del tiempo de cálculo necesario para su convergencia, o lo que es lo mismo, el número de iteraciones necesarias para alcanzarla. Esta cuestión depende de la estructura y características específicas de los problemas a los que se aplique.

Recuerdese que en los métodos duales, a diferencia de los primales, las soluciones sólo son factibles cuando la convergencia se ha alcanzado.

Dos últimas observaciones nos conducirán al tema del apartado siguiente :

- la introducción de restricciones instantáneas no es imposible teóricamente. Podría introducirse una variable de holgura por cada una de las restricciones y aumentar correspondientemente los vectores X, P_x etc. utilizados. Sin embargo, un exceso de tal tipo de restricciones puede introducir dimensiones tales que el método deje de ser aplicable.
- observese que la aplicación del método no pasa por ningún procedimiento de descomposición como los que generan métodos duales puros tipo Lasdon-Tamura. Ello es debido a que el problema modificado tiene elementos de la forma $X(k+1) f(x(k), u(k))$ que inducen una encadenación de lo que hubieran sido problemas locales como los que aparecen en el apartado siguiente, en todo el cual suponemos verificada la hipótesis de convexidad local.

Por la hipótesis enunciada, la Teoría de la dualidad local es aplicable sin precisar de la modificación de Hestenes. En tal caso, la aplicación de los métodos duales a problemas de control óptimo conduce a una muy eficaz descomposición de las operaciones del 1^{er} nivel (evaluación de la función dual), en una serie de problemas locales independientes. Cada uno de ellos corresponde a uno de los intervalos, K , discretizados del tiempo y en él solo intervienen las variables de estado y control de dicho intervalo, es decir $x(k)$, $U(k)$.

Este resultado es la extensión a problemas discretos por Tamura (T3) del análisis de problemas continuos de control óptimo de Pearson (P4) y Lasdon(L5). El método es tan importante conceptual como computacionalmente. Está en la base de todos los procedimientos de descentralización-descomposición del capítulo V y genera una información dual a la que es aplicable directamente todo el aparato interpretativo del capítulo II.

Características básicas del desarrollo expuesto a continuación son:

- a) planteamiento del problema en dualidad parcial, afectando solamente a las restricciones igualdad representadas por las ecuaciones de estado.
- b) aplicado a sistemas con efectos retardados sin necesidad de aumentar el espacio de estado y de control (técnica habitualmente empleada).
- c) utilización del método del gradiente reducido generalizado para resolver los problemas locales del 1^{er} nivel bajo sus restricciones desigualdad correspondientes.

Como en todo algoritmo dual, las soluciones solo son

factibles en la convergencia. De aquí la denominación de "infeasible methods" con los que se les denomina en las referencias (L5,W1,P4,S1) donde se expone detalladamente su justificación teórica.

Supongamos de nuevo el problema:

$$\text{Min } J(x) = \text{Min}_{\mu} \left[\Phi(x(N)) + \sum_{k=0}^{N-1} f_0(x(k), u(k), k) \right]$$

con las ecuaciones de estado presentando retrasos en el tiempo de la forma:

$$\begin{aligned} x(k+1) &= f(x(k), u(k), k) + f(x(k-1), u(k-1), k-1) + \dots + \\ &+ f(x(k-0), u(k-0), k-0) = \\ &= \sum_{i=0}^0 f(x(k-i), u(k-i), k-i) \end{aligned}$$

y con las trayectorias restringidas a discurrir por las zonas del espacio de control-estado definidas por las restricciones instantáneas :

$$R(x(k), u(k), k) \leq 0$$

$$k=0, N-1$$

con el estado inicial supuesto fijo

$$x(0) = x_0$$

y el estado final de la trayectoria perteneciente a un cierto dominio definido por las restricciones :

$$R(x(N)) \leq 0$$

Llamando $\Psi(k+1)$ a la variable adjunta (multiplicadores asociados a las restricciones igualdad) correspondiente a la ecuación de estado:

$$\sum_{i=0}^0 f(x(k-i), u(k-i), k-i) - x(k+1) = 0$$

y formando el Lagrangiano del problema (de acuerdo con

a))

$$\begin{aligned} \mathcal{L}(x, u, \psi) &= \Phi(x(N)) + \sum_{k=0}^{N-1} f_0(x(k), u(k), k) + \\ &+ \sum_{k=0}^{N-1} \psi'(k+1) \left[\sum_{i=0}^{\infty} f(x(k-i), u(k-i), k-i) - x(k+1) \right] \\ &= \Phi(x(N)) + \\ &+ \sum_{k=0}^{N-1} \left[f_0(x(k), u(k), k) + \psi'(k+1) \left[\sum_{i=0}^{\infty} f(x(k-i), u(k-i), k-i) - x(k+1) \right] \right] \end{aligned}$$

Para tener en cuenta la existencia de retrasos en el tiempo (dependencia de $x(k+1)$ en valores de estados y controles anteriores a k), modificamos la definición de Hamiltoniano dada en el Apéndice II, donde tales retrasos no existían.

Desarrollando el segundo término del sumatorio sobre k quedaría, teniendo en cuenta que los valores de variables con índices negativos se consideran nulos:

$$\begin{aligned} k=0 & \quad \psi(1) [f(x(0), u(0), 0) - x(1)] \\ k=1 & \quad \psi(2) [f(x(1), u(1), 1) + f(x(0), u(0), 0) - x(2)] \\ k=2 & \quad \psi(3) [f(x(2), u(2), 2) + f(x(1), u(1), 1) + f(x(0), u(0), 0) \\ & \quad \quad - x(3)] \\ & \quad \vdots \\ k & \quad \psi(k+1) [f(x(k), u(k), k) + \dots + f(x(k-\theta), u(k-\theta), 0) \\ & \quad \quad - x(k+1)] \\ & \quad \vdots \\ N-1 & \quad \psi(N) [f(x(N-1), u(N-1), N-1) + \dots + f(x(N-1-\theta), \\ & \quad \quad \quad u(N-1-\theta), N-1-\theta) - x(N)] \end{aligned}$$

Agrupando los valores de $\psi(i)$ que afectan a una misma

$$\sum_{k=0}^{N-1} f(x(k), u(k), k) \sum_{i=1}^{Q+1} \psi'(k+i) - \sum_{k=0}^{N-1} \psi'(k+1) x(k+1)$$

Definiendo la función Hamiltoniano como :

$$H[x(k), u(k), \psi, k] = f_0(x(k), u(k), k) + f(x(k), u(k), k) \sum_{i=1}^{Q+1} \psi'(k+i)$$

el Langrangiano puede escribirse :

$$\begin{aligned} \mathcal{L}(x, u, \psi) &= \Phi(x(N)) + \sum_{k=0}^{N-1} H[x(k), u(k), \psi, k] - \\ &\quad - \sum_{k=0}^{N-1} \psi'(k+1) x(k+1) = \\ &= \Phi(x(N)) - \psi'(N) x(N) + \sum_{k=0, N-1} [H[x(k), u(k), \psi, k] - \psi'(k) x(k)] \end{aligned}$$

con la convención: $\psi(0)=0$

La importancia de tales definiciones y agrupaciones es que, ahora, las únicas variables que aparecen en cada término del sumatorio, son las de estado y control del instante correspondiente

La evaluación de la función dual, $\min_{x, u} \mathcal{L}(x, u, \psi)$, bajo el conjunto de restricciones instantáneas, es la tarea encomendada al primer nivel del método dual. Para así hacerlo, este recibe unos ciertos valores de las variables duales ψ fijadas por el 2º nivel (arbitrarias en la 1ª iteración). Una vez que estos valores son introducidos en $\mathcal{L}(x, u, \psi)$, el problema del primer nivel queda naturalmente descompuesto en $N+1$ problemas independientes, cada uno de ellos sometido a sus propias restricciones desigualdad.

El número de variables que aparecen en cada uno de los problemas locales es $n+m$, $(x(k), u(k))$

$$K = N$$

$$\min_{x(N)} \left\{ \Phi(x(N)) - \Psi'(N)x(N) \right\}$$

$$\text{con: } R(x(N)) \leq 0$$

⋮

⋮

$$K$$

$$\min_{x(k), u(k)} \left\{ H[x(k), u(k), \Psi] - \Psi'(k)x(k) \right\}$$

$$\text{con: } R(x(k), u(k), k) \leq 0$$

es decir:

$$\min_{x(k), u(k)} \left\{ f_0(x(k), u(k), k) + f(x(k), u(k), k) \sum_{i=1}^0 \Psi'(k+i) - \Psi'(k)x(k) \right\}$$

$$\text{con: } R(x(k), u(k), k) \leq 0$$

⋮

⋮

$$K = 0$$

$$\min_{u(0)} \left\{ f_0(x(0), u(0), 0) + f(x(0), u(0), 0) \sum_{i=1}^0 \Psi'(i) \right\}$$

$$\text{con } x(0) = x_0$$

$$R(x(0), u(0), 0) \leq 0$$

A pesar de la existencia de los retrasos en el tiempo, la evaluación de la función dual se reduce a la resolución de $N + 1$ programas matemáticos independientes. Cada uno de ellos corresponde a uno de los instantes de tiempo $k = 0, N$ y se expresa únicamente en función de las variables de estado y control $(x(k), u(k))$ correspondientes a dicho instante. Sus funciones objetivo se construyen con los Hamiltonianos mediante las variables duales y están sometidos a las correspondientes restricciones instantáneas sobre la trayectoria.

La resolución de cada uno de estos problemas locales del primer nivel puede realizarse mediante el referido método del gradiente reducido generalizado.

Se substituye así un programa con $N(n+m)$ variables por $N + 1$ programas : $N - 1$ de ellos con $n + m$ variables, uno (el correspondiente al estado inicial) con m variables (las $u(0)$) y el que completa el cómputo (correspondiente a $k = N$) con las n variables $x(N)$. El lector habrá observado que los retrasos en el tiempo tampoco exigen ampliar las dimensiones del vector dual. Esta observación tiene interés cara a las modificaciones en los $\psi(k)$ efectuadas por el primer nivel. Este recibe los resultados de los subproblemas locales, es decir, una evaluación de la función dual :

$$\Phi(\psi) = \min_{x, u} L(x, u, \psi)$$

y debe de maximizarla con respecto a ψ . El problema del segundo nivel es pues un programa matemático sin restricciones (puesto que $\psi(k)$ son multiplicadores asociados a restricciones igualdad) con un total de $N.n$ variables (las $\psi(k), k=1, N$). En el 2º nivel no existe pues descomposición del problema. Pero este es mucho más fácil de resolver dada la ausencia de restricciones, la conca-

vidad de la función dual y el resultado citado acerca de su gradiente :

$$\nabla_{\psi(k)} \Phi(\psi) = \sum_{i=0}^0 f(x(k-i), u(k-i), k-i) - x(k+1)$$

donde los valores de $x(k+1), x(k), \dots, x(k-0), u(k), \dots, u(k-0)$ son los calculados en el primer nivel.

Observe que los componentes del gradiente de la función dual son los errores cometidos en la satisfacción de las restricciones, lo cual es realmente ilustrativo en su interpretación económica.

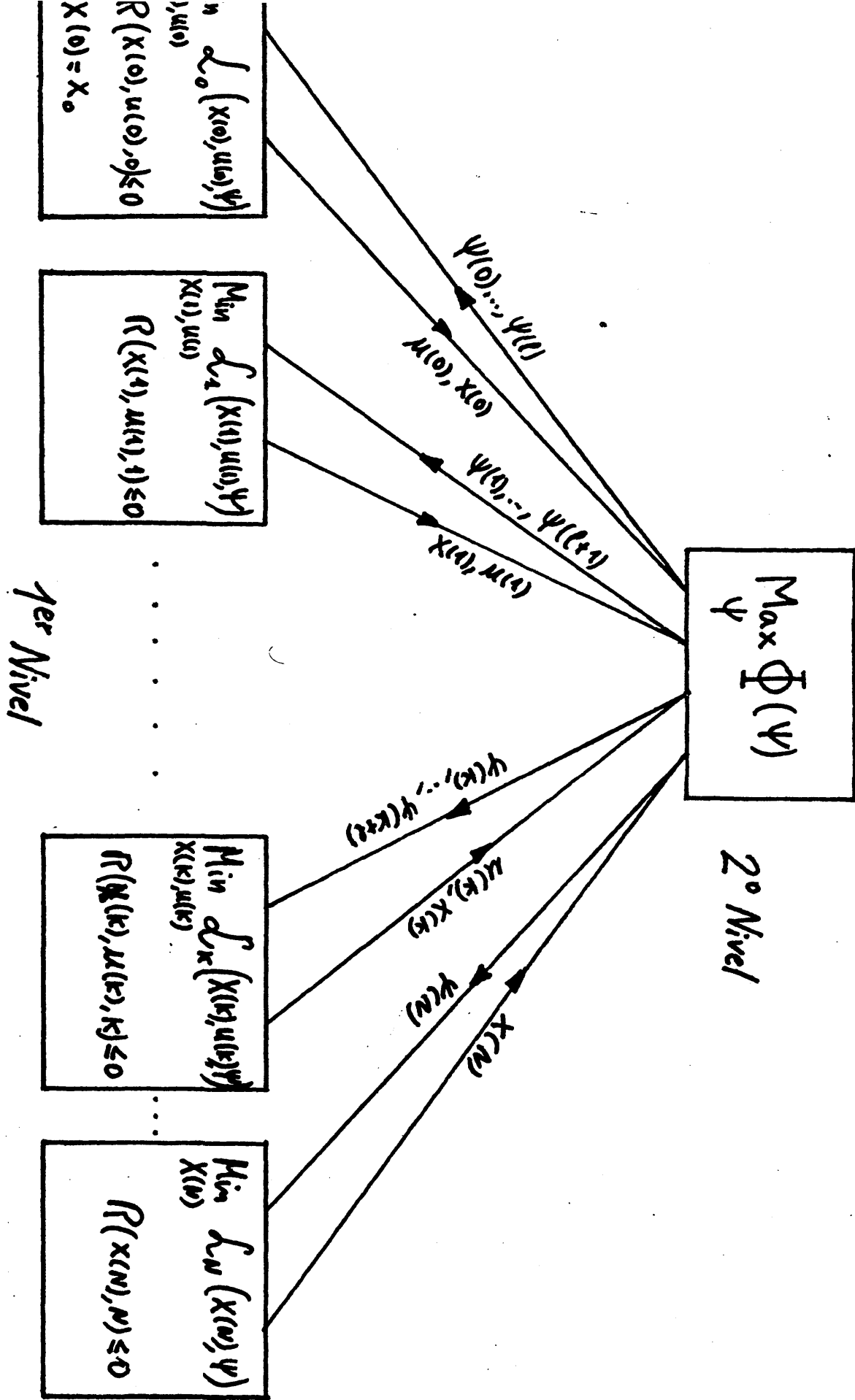
La interacción entre niveles y la transmisión sucesiva de información primal y dual proseguirá de acuerdo con el esquema de la figura FIII.8

Si la hipótesis de convexidad local se satisface, el método convergerá hacia un punto solución en el que coincidirán los valores de los problemas primal y dual :

$$\max_{\lambda} \Phi(\lambda) = \min_{\mu} J(x, u)$$

Observe que este sólo será un óptimo local si el programa no es convexo. (Lasdon (15)) . Observe también que la reducción en la carga computacional, con ser enorme, no deja que el problema sea realmente difícil y exigente. En efecto, los $(N + 1)$ programas con $(n + m)$ variables del 1^{er} nivel y el programa sin restricciones con $N.n$ variables del 2^o deben resolverse tantas veces como iteraciones se realicen antes de alcanzar la convergencia. Si el número de estas es de NI, hemos de resolver un total de $NI(N + 2)$ programas matemáticos.

Todo ello con dos agravantes que no deben ser olvidados:



- Si el proceso es abandonado antes de conseguir su convergencia, el esfuerzo realizado hasta entonces es esteril, en cuanto a la solución del problema específicamente planteado se refiere, puesto que el resultado de la última iteración corresponde a un punto no factible. En cambio el método primal propuesto genera una secuencia de puntos, cada uno representando una mejora sobre el anterior, y siendo todos ellos factibles. Sin embargo, como analizaremos en el capítulo V, la información generada puede ser de gran utilidad cuando las restricciones del problema no se conocen con gran precisión.
- Si la hipótesis de convexidad local no se verificaba, el punto hacia el que convergeremos no tiene porque corresponder ni siquiera a un óptimo local. Ello podrá ser testado fácilmente comparando los valores de la función dual y primal en el punto solución.
Importa señalar aquí que la no convexidad del problema no excluye la posibilidad de aplicación de los métodos duales, aunque deje de estar asegurada su eficacia.

La idea de descomposición asociada a métodos duales la retenemos para un más profundo desarrollo en el capítulo V, en el que introduciremos varios centros de decisión asociados con los subsistemas interrelacionados, que componen la estructura del problema global.

En el capítulo siguiente, analizamos la aplicación de los métodos aquí expuestos a problemas de planificación multisectoriales, como una alternativa a los modelos lineal cuadráticos habitualmente empleados.

CAPITULO IV

CONTROL OPTIMO Y PLANIFICACION DINAMICA MULTISECTORIAL CON CENTRO DE DECISION UNICO

" A smaller plan will present less
problems, but will solve none".

Pitamber Pant. (P2)

IV. 1 PRESENTACION

IV. 2 EL MODELO LINEAL-CUADRATICO.

IV. 2. 1 REGLAS DE DECISION LINEALES Y "FEED-BACK"
CONTROL.

IV. 2. 2 CARACTERISTICAS DEL SISTEMA LINEAL.

IV. 2. 3 CARÁCTERISTICAS DE LAS FUNCIONES OBJETIVO
CUADRATICAS.

IV. 3 MODELO DE PLANIFICACION DINAMICA MULTISECTORIAL NO LINEAL.

IV. 3. 1 PLANTEAMIENTO DEL MODELO.

IV. 3. 2 FORMULACION DEL PROBLEMA DE CONTROL.

IV. 3. 3 RESOLUCION POR EL METODO DUAL DE LOS MULTI
PLICADORES.

IV. 4 LOS LIMITES DE UNA SOLUCION GLOBAL.

Los modelos matemáticos de sistemas económicos nacionales desarrollados hasta la fecha, pueden ser divididos en dos grandes grupos: Los modelos a corto plazo de política monetaria y fiscal y los modelos a largo plazo de programación de la inversión y análisis de estrategias de desarrollo.

Según la metodología matemática aplicada, ambos tipos pueden clasificarse en modelos de simulación y de programación. El objetivo de los pertenecientes al primero de estos grupos es el de analizar los efectos de cambios en las políticas monetaria y fiscal en los niveles de precios, producción, empleo, inversión, etc... en un periodo que se extiende entre 3 meses y 2 años después de la aplicación de tales políticas. La amplitud de los periodos de tiempo considerados en la construcción del modelo es corto, normalmente mensual o trimestral.

En contrapartida, los modelos que analizan las estrategias de desarrollo a largo plazo pretenden fundamentalmente analizar las consecuencias de las distintas alternativas de asignación en el tiempo de la inversión en los distintos sectores de la economía, en las trayectorias de producción y/o consumo en un periodo de 5 a 20 años. El periodo de tiempo utilizado es generalmente el anual.

El objetivo de hacer explícita esta diferenciación es el de señalar que nuestro interés lo constituyen aquí los modelos del segundo tipo.

Desde un punto de vista de teoría del control podemos interpretar tal división diciendo que la segunda clase de modelos pretende determinar la trayectoria óptima (en el sentido dado por una funcional objetivo definida con y por criterios políticos) del sistema económico, mientras que

la función de modelos del primer tipo es el de mantener a la economía, en las distintas etapas de su devenir temporal, en las cercanías de la trayectoria a largo plazo así determinada, Es decir, cumplen fundamentalmente una función de regulación.

La evolución de los métodos matemáticos empleados en modelos del primer tipo viene a corroborar tales consideraciones. En efecto, desde que las técnicas de control óptimo han venido a substituir a las de simulación en la evaluación sobre un modelo econométrico de las medidas de política económica, los modelos desarrollados han sido casi exclusivamente los lineal-cuadráticos.

Sabido es que los modelos lineal-cuadráticos son de larga tradición, tanto en ingeniería como en economía, y que se les conoce con el nombre genérico de "reguladores".

Muchas de sus bases teóricas han sido aportaciones, precursoras e ignoradas, de economistas, en cuyo campo han dado origen a las famosas reglas de decisión lineales (Theil, (T6)). Las razones de la renovada atención de que han sido objeto en sus aplicaciones de política económica (Pindyck (P15), Friedman (F8), Norman (N3), Turnovsky (T9), Chow (C7)) las encontramos en su relación con la teoría del multiplicador, el intuitivo funcionamiento de su sistema de valoración y a sus conocidas propiedades computacionales :

- aplicación del principio de certeza-equivalencia, o teorema de separación, para tener en cuenta perturbaciones aleatorias. (Apéndice IV)
- control óptimo expresado como una función lineal del estado del sistema, que ha institucionalizado universalmente la terminología anglosajona del " feed-back".
- cálculo del control reducido a la solución de un sis-

(Bryson (B5)).

Creemos sin embargo que las limitaciones implícitas en tal tipo de formulación no aconsejan su extensión a modelos de análisis a largo plazo de las estrategias del desarrollo. Creemos también que el empleo de nuevos métodos como los expuestos en el capítulo III, debe poder desplazar el equilibrio "facilidad de resolución-realismo en la formulación", a favor de modelos más generales.

El punto IV. 2 expone las razones que sustentan tal toma de posición. Dado el extenso tratamiento de que ha sido objeto el modelo lineal-cuadrático, nos limitaremos a analizar la intrínseca similitud entre las reglas de decisión lineales y la solución de problemas de regulación por métodos variacionales y a ponderar sus ventajas computacionales frente a sus limitaciones.

Especial atención es dedicada al componente sistema lineal de estado. Su importancia trasciende al conjunto del modelo puesto que ha constituido, junto con las técnicas de programación lineal, el método e instrumento favoritos y cuasi exclusivos de planificación económica hasta principio de la actual década.

Del punto IV. 2 obtenemos una serie de conclusiones. Entre ellas figuran la indeseabilidad de relaciones lineales en planteamientos a largo plazo. La necesidad de complementar, con restricciones desigualdad en las variables de estado, las ecuaciones de evolución lineales que aparecen al utilizar las fuentes de datos que constituyen los sistemas dinámicos input-output. La importancia de traducir las preferencias políticas en la forma de una función objetivo que no precise una trayectoria "a priori" para ser evaluada y completarla mediante la toma en con-

consideración de restricciones sobre el estado final. El interés, en ocasiones, de considerar problemas de mínimo tiempo en vez de un objetivo construido por una funcional de las variables de estado y/o control.

Tales conclusiones orientan el restante contenido del capítulo. En el punto IV. 3 desarrollamos un modelo de planificación multisectorial dinámico no lineal del tipo Kendrick (K_6). Aplicando los algoritmos del capítulo III podemos extender el modelo a más sectores que los originalmente considerados por Kendrick, al tiempo que reforzamos la introducción de preferencias de tipo político mediante la consideración de restricciones instantaneas sobre las variables.

IV. 2 EL MODELO LINEAL-CUADRATICO.

IV. 2.1 REGLAS DE DECISION LINEALES Y "FEED-BACK" CONTROL.

El desarrollo de los modelos lineal-cuadráticos constituye un flagrante ejemplo de incomunicación intelectual entre disciplinas. Podemos resumir su trayectoria histórica en los siguientes puntos:

En 1956-'57, Simon (S_4) y Theil (T_5) exponen por primera vez el principio de certeza-equivalencia en dos artículos, hoy celebres, de "Econometría". Cinco años tuvieron que transcurrir antes de que fuera redescubierto por Tou (T_7) y Gunckel-Franklin (C_6) para ser utilizado en sistemas lineales de control.

En 1962 Holt (H_7) expone las reglas de decisión lineales como una extensión del concepto del multiplicador. Theil (T_6) las generaliza y extiende al análisis de modelos lineales dinámicos.

La contribución de Phillips, contemporánea de los primeros trabajos de Simon y Theil, puso de relieve la importancia de los retrasos en el tiempo ("lags") y las consecuencias de desfases en la aplicación de medidas de política económica.

Sus trabajos, como tantos otros, durmieron el sueño de lo teórico hasta que el desarrollo del ordenador digital y los algoritmos numéricos, posibilitaron el reciente resurgir de sus aplicaciones económicas. Pindyck (P15) es, quizá, el autor de la más conocida de las mismas.

Su método es una extensión del problema del regulador. Pindyck usa una función objetivo cuadrática para hacer que la solución obtenida siga, lo mas cerca posible, la trayectoria temporal previamente especificada de las variables de estado $\hat{x}(t)$ y de control $\hat{u}(t)$. Las desviaciones son penalizadas cuadráticamente y ponderadas relativamente por unas matrices Q,R, también predeterminadas, que traducen preferencias políticas.

Analíticamente, el problema se expresa, en régimen discreto :

Hallar

$$x^*(k), u^*(k), k=0, \dots, N$$

tal que :

$$x^*(k+1) = x^*(k) + A x^*(k) + B u^*(k) + C z(k)$$

$$x^*(0) = x_0$$

$$(k, = 0, N-1)$$

y que minimice :

$$J = \frac{1}{2} \sum \|x(k) - \hat{x}(k)\|_Q^2 + \|u(k) - \hat{u}(k)\|_R^2$$

donde :

A, B, C : matrices conocidas de las ecuaciones de estado.

X_0 : condiciones iniciales conocidas.

$Z(k)$: vector de variables exógenas.

Friedman (F8) extiende el modelo considerando una función cuadrática por intervalos y determinando endogenamente el horizonte del plan. Corrige así algunas de las deficiencias del modelo lineal-cuadrático.

Puesto que uno de los requisitos básicos de un sistema lineal-cuadrático es la disponibilidad de un modelo lineal, el desarrollo de los sistemas dinámicos de input-output ha posibilitado muchas de sus aplicaciones macroeconómicas.

Livesey (L8) y Brody (B7), basándose en trabajos de Dorfman (D11), Stone (S11) y Jorgenson (J3), han analizado las dificultades que encierra la adopción de un sistema input-output dinámico como sistema de ecuaciones de estado. Básicamente son debidas a la atribución de caracter causal a lo que en principio es una identidad contable que refleja la hipótesis de plena utilización del stock de capital.

Las relaciones lineales usadas por Theil sí representan relaciones causales, pero los modelos formulados lo son en la forma reducida. Ello diferencia el proceso de cálculo de las soluciones, de aquel que se obtendría mediante la aplicación directa de los métodos de control óptimo, puesto que estos expresan sus ecuaciones en la forma del "espacio de estado" ("state space form").

Esta es la forma en la que se expresan los modelos input-output dinámicos y ha sido la utilizada en casi todas las recientes aplicaciones económicas del método lineal-cuadrático. Obsérvese que ello implica el abandono de la forma reducida en los modelos econométricos utilizados.

El análisis de las dos cuestiones así planteadas :

- a) relaciones y diferencias entre las reglas de decisión lineales por una parte y la solución que la teoría del control óptimo facilita de los problemas lineal-cuadráticos por otra.
- b) representación de las ecuaciones de estado mediante un sistema dinámico input-output.

Es conducido por dos factores determinantes :

- a) la estructura formal que se adopte para las ecuaciones del sistema.
- b) las consecuencias que se deriven de la causalidad que tales relaciones deben necesariamente representar.

Observe que la primera cuestión se refiere al proceso de optimización del sistema una vez que ha sido resuelta la descripción de su dinámica. En cambio, la segunda de ellas no hace intervenir ninguna consideración relativa a la optimización del sistema que trata de describir.

Dedicando el punto siguiente a este último tema, nos proponemos ahora presentar las reglas de decisión lineales de Theil como un caso particular de la ley general de control lineal en realimentación (feed-back). En lo fundamental, la exposición está basada en Livesey (L8).

La derivación de reglas de decisión lineales está ya contenida, esencialmente, en el esquemático y estático modelo del multiplicador :

$$\left. \begin{array}{l} Y = C + I + G \\ C = cY \\ c < 1 \end{array} \right\} \Rightarrow Y = \frac{I + G}{1 - c} \quad (1)$$

al que Holt (H7) añade una función objetivo cuadrática:

$$\text{Min}_G J = (Y - Y^*)^2 + a(G - G^*)^2 + b(Y - Y^*)(G - G^*)$$

en la que Y^*, G^* son los niveles deseados de demanda global y gasto público. El último término de la función objetivo expresa la hipótesis de que las desviaciones del mismo signo en las variables Y, G refuerzan su efecto negativo.

La derivada del objetivo cuadrático producirá una relación lineal :

$$\frac{\partial J}{\partial G} = 0 = b(Y - Y^*) + 2a(G - G^*)$$

que junto con la ecuación de estado (1) produce la llamada "regla de decisión lineal" (linear decision rule):

$$G = \frac{1}{2a + \frac{b}{1-c}} \left[2aG^* + Y^* - \frac{bI}{1-c} \right] \quad (1)$$

Theil considera la maximización de una función cuadrática:

$$\text{Max } J(x, y) = \frac{1}{2} [x' B x + y' C y] \quad (2)$$

sometida a las restricciones representadas por un modelo econométrico lineal dinámico expresado en la forma reducida :

$$y = Ax + S \quad (3)$$

en donde X es el m -vector de los instrumentos, y es el n -vector de las variables objetivo y S es un vector de variables exógenas (constantes).

Los vectores X, y contienen las variables correspondientes a cada uno de los periodos en el interior del horizonte del plan T :

$$X = \begin{bmatrix} X(1) \\ \vdots \\ X(k) \\ \vdots \\ X(T) \end{bmatrix} = \begin{bmatrix} X_1(1) \\ \vdots \\ X_m(1) \\ \vdots \\ X_1(k) \\ \vdots \\ X_m(k) \\ \vdots \\ X_1(T) \\ \vdots \\ X_m(T) \end{bmatrix}$$

$$Y = \begin{bmatrix} Y(1) \\ \vdots \\ Y(k) \\ \vdots \\ Y(T) \end{bmatrix} = \begin{bmatrix} Y_1(1) \\ \vdots \\ Y_n(1) \\ \vdots \\ Y_1(k) \\ \vdots \\ Y_n(k) \\ \vdots \\ Y_1(T) \\ \vdots \\ Y_n(T) \end{bmatrix}$$

La matriz A es una matriz triangular inferior compuesta por $\frac{T(T-1)}{2} + T$ submatrices A_{ij} no idénticamente nulas, de dimensiones $n \times m$. Todas las submatrices A_{ij} , $j > i$ son idénticamente nulas :

$$A_{ij} \equiv 0 \quad \forall j > i$$

Los elementos $(a_{\kappa\epsilon})_{ij}$ de las submatrices no nulas A_{mn} representan la eficacia de la variable instrumental $X_\epsilon(j)$ sobre la variable objetivo $Y_\kappa(i)$.

Puesto que ningún instrumento puede causar efectos en periodos anteriores a la fecha de su aplicación, es obvio que la matriz A debe presentar la forma de la figura:

$$\begin{bmatrix} A_{11} & & & & & \\ & A_{21} & A_{22} & & & \\ & \vdots & \vdots & \ddots & & \\ & & \vdots & \ddots & & \\ & & & & & 0 \\ & A_{k1} & A_{k2} & \dots & A_{kk} & \\ & \vdots & \vdots & \vdots & \vdots & \\ & & \vdots & \vdots & \vdots & \\ & A_{T1} & A_{T2} & \dots & A_{TK} & \dots & A_{TT} \end{bmatrix}$$

Substituyendo las ecuaciones del modelo (3), en la función objetivo (2) se obtiene una expresión cuadrática en x :

$$\begin{aligned} J(x, Ax+s) &= \frac{1}{2} [x' B x + (x' A' + s') C (Ax + s)] = \\ &= K_0 + K_1 x + \frac{1}{2} x' K_2 x \end{aligned}$$

Expresión en la que :

$$K_0 = \frac{1}{2} s' C s$$

$$K_1 = s' C A$$

$$K_2 = B + A' C A$$

Con la hipótesis de que K_2 es regular y definida negativa, el valor óptimo del vector de instrumentos viene dado por :

$$\frac{\partial J}{\partial x} = K_1 + K_2 x = 0$$

$$x^* = -K_2^{-1} K_1' = -[B + A'CA]^{-1} [A'C]s$$

Mediante la correspondiente substitución de s , se llega a la clásica ley lineal que expresa los valores óptimos de los instrumentos mediante una función lineal de las variables objetivo :

$$x^* = -B^{-1} [C + A'C] y$$

En Teoria del control llamaríamos variables de estado y de control a $y(t)$, $x(t)$ respectivamente. Abandonando la forma reducida, las ecuaciones del sistema se escriben, en régimen continuo :

$$\dot{y}(t) = F(t)y(t) + G(t)x(t)$$

o, en discreto :

$$y(t+1) = F(t)y(t) + G(t)x(t) \\ (t = 0, T-1)$$

La estructura de la solución obtenida por la Teoria del Control Óptimo es la misma que en las reglas de decisión lineales. Tanto en el caso continuo como en el discreto, el control óptimo se obtiene como una función lineal del estado del sistema. Tal ley de control viene dada por (Bryson, (B5)) :

a) Caso continuo :

$$\text{Max } \frac{1}{2} \int_0^T \left[y'(t) C(t) y(t) + x'(t) B(t) x(t) \right] dt$$

$$\dot{y}(t) = F(t) y(t) + G(t) x(t)$$

$$y(0) = y_0$$

Solución:

$$x^*(t) = -[B(t)]^{-1} A'(t) S(t) y(t)$$

en donde $S(t)$ es la solución del sistema de ecuaciones diferenciales tipo Riccati:

$$\dot{S} = -SF - F'S + SAB^{-1}A'S - C$$

$$S(T) = 0$$

b) Caso discreto (Kirk, (K10)) :

$$\text{Max } \frac{1}{2} \sum_{k=0}^N \left[y'(k) C(k) y(k) + x'(k) B(k) x(k) \right]$$

$$y(k+1) = F(k) y(k) + G(k) x(k) \\ (k=0, N-1)$$

$$y(0) = y_0$$

Solución:

$$x^*(N-k) = -H(N-k) y(N-k) \\ (k=0, N)$$

donde :

$$H(N-k) = \pi_{1k}^{-1} \pi_{2k}$$

$$\pi_{1k} = [B(N-k) + G'(N-k) P(k-1) G(N-k)]^{-1}$$

$$\pi_{2k} = G'(N-k) P(k-1) F(N-k)$$

y

$$P(k) = [F(N-k) + G(N-k) H(N-k)]^{-1} \cdot$$

$$\cdot P(k-1) \cdot$$

$$\cdot [F(N-k) + G(N-k) H(N-k)] +$$

$$+ H'(N-k) B(N-k) H(N-k) + C(N-k)$$

$$(k=0, N)$$

con $p(-1) = 0$

La solución se obtiene, como indican las anteriores expresiones, mediante una evaluación secuencial alternativa de las matrices H , y P a partir de $p(-1) = 0$.

El planteamiento lineal cuadrático es el único que posee la propiedad de que el valor óptimo de los instrumentos (controles), se exprese como una función lineal de los objetivos (estados). Las reglas de decisión de Theil aparecen como un caso particular de la ley general de control lineal en realización ("feed-back").

La idéntica estructura de las soluciones no se extiende a los respectivos procedimientos de cálculo. Ello es debido a que, en ambas formulaciones, se han utilizado métodos radicalmente distintos para describir la dinámica del sistema.

La formulación de Theil no es la mas práctica para representar un sistema dinámico, y los resultados a que conduce, aunque analíticamente concisos, no son necesariamente eficientes como métodos de solución numérica. En particular la

inversión de la matriz K_2 puede presentar serias dificultades, dadas sus dimensiones. La formulación en el espacio de estados, reduce tal dificultad al disminuir la dimensión de las matrices empleadas.

El análisis de las diferencias entre ambos procedimientos de cálculo, debidas al diferente significado de las matrices A, F, G puede avanzar un paso más estudiando la relación que existe entre tales matrices.

Considerando los dos sistemas de representación del sistema dinámico :

a) forma reducida :

$$y = Ax + s$$

b) formulación en el espacio de estados :

$$\begin{aligned} y(t+1) &= F(t)y(t) + G(t)x(t) \\ y(0) &= y_0 \end{aligned} \quad (4) \quad (t=0, T-1)$$

se puede mostrar que la matriz A está constituida fundamentalmente por las matrices de transición del sistema de ecuaciones en diferencias finitas (4) utilizado en la formulación b). En efecto :

Por aplicación reiterada de (4), se obtiene :

$$y(t+k) = \Phi(t, t+k)y(t) + \sum_{i=1}^k \Phi(t+i, t+k)G(t+i-1)x(t+i-1) \quad (5)$$

donde $\Phi(t, t+k)$ es la matriz de transición, dada por :

$$\Phi(t, t+k) = \prod_{i=t}^{t+k-1} F(i)$$

$$\Phi(t, t) = I$$

Escribiendo la ecuación (5) para $t=0$ y los sucesivos valores crecientes de K , $K=1, 2 \dots T$, obtenemos :

$$y(1) = G(0)x(0) + \Phi(0,1)y(0)$$

$$y(2) = \Phi(1,2)G(0)x(0) + G(1)x(1) + \Phi(0,2)y(0)$$

...

$$y(k) = \Phi(1,k)G(0)x(0) + \Phi(2,k)G(1)x(1) + \dots \\ + G(k-1)x(k-1) + \Phi(0,k)y(0)$$

...

$$y(T) = \Phi(1,T)G(0)x(0) + \Phi(2,T)G(1)x(1) + \dots \\ + G(T-1)x(T-1) + \Phi(0,T)y(0)$$

Matricialmente:

$$y = Ax + s \quad (6)$$

donde:

$$A = \begin{bmatrix} G(0) & 0 & & & \\ \Phi(1,2)G(0) & G(1) & & & \\ \vdots & \vdots & \ddots & & \\ \vdots & \vdots & \ddots & 0 & \\ \Phi(1,k)G(0) & \Phi(2,k)G(1) & \dots & G(k-1) & \dots \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \vdots & \ddots & 0 \\ \Phi(1,T)G(0) & \Phi(2,T)G(1) & \dots & \Phi(k,T)G(k-1) & \dots & G(T-1) \end{bmatrix}$$

$$Y = \begin{bmatrix} Y(1) \\ Y(2) \\ \vdots \\ Y(k) \\ \vdots \\ Y(\tau) \end{bmatrix} \quad X = \begin{bmatrix} X(0) \\ X(1) \\ \vdots \\ X(k) \\ \vdots \\ X(\tau-1) \end{bmatrix} \quad S = \begin{bmatrix} \Phi(0,1) \\ \vdots \\ \Phi(0,k) \\ \vdots \\ \Phi(0,\tau) \end{bmatrix} Y(0)$$

Aparece así la matriz A expresada en función de las matrices de transición del sistema dinámico (4). La expresión (6) tiene la virtud de mostrar claramente la enorme talla de la matriz A , en comparación con la de las matrices F , G utilizadas en las ecuaciones (4).

Por ello, el cálculo de los valores óptimos de las variables instrumento según las reglas lineales de Theil representa, a pesar de su aparente sencillez, un colosal problema de cálculo. Por el contrario, la aplicación de las reglas de cálculo secuencial de los controles dadas por el principio del máximo, tiene una apariencia más compleja pero representa un procedimiento infinitamente más fácil de desarrollar en la práctica.

Ello no podría ser de otro modo. La derivación de las reglas de decisión lineales no hace intervenir ningún procedimiento de optimización específicamente dinámico. Una vez aplicadas (suponiendo superables los obstáculos que su sencilla terminología encierra) obtenemos en bloque toda la trayectoria temporal de las variables instrumentales (controles). En cambio, la aplicación del principio del máximo (o del principio de optimalidad), (específicos instrumentos dinámicos), permite descentralizar temporal-

mente el cálculo de los controles. Este se lleva a cabo de una manera secuencial en el tiempo sin dejar por ello de garantizar el óptimo global.

En consecuencia, el cálculo puede organizarse de forma mucho mas ventajosa.

En resumen: las reglas de decisión lineales permiten obtener una solución cuya estructura es la misma que la que se obtiene aplicando el principio del máximo a un problema lineal cuadrático. Los métodos de cálculo son muy diferentes porque lo son las dos formas de representar la dinámica del sistema. El propuesto por la Teoria del Control óptimo es mucho mas eficiente.

Evidentemente, el interés de las reglas de decisión lineales derivadas a partir de un modelo lineal cuadrático no esta en la forma de cálculo propuesta, sino en ser esta la formulación con la que Theil demostró el teorema de certeza-equivalencia, suponiendo determinadas hipótesis en el comportamiento de las variables exógenas s , ahora supuestas variables aleatorias. El Anexo IV recoge la formulación del teorema de certeza-equivalencia en la forma definitivamente adquirida en teoria del control óptimo. A el nos referimos de nuevo al ponderar las ventajas e inconvenientes del uso de modelos lineal-cuadráticos.

IV. 2. 2. CARACTERISTICAS DEL SISTEMA LINEAL.

Las referencias básicas de este apartado son los artículos de Livesey (L8) y la obra de Brody (B6), así como los autores por ellos citados.

Livesey analiza los distintos modelos econométricos en función de la linealidad que les pueda permitir ser utilizados como ecuaciones de estado de un sistema lineal-cuadrático. Su conclusión es que " en general los modelos econométricos no son lineales y que cuando un autor precisa uno de tal característica tiene que construirlo especialmente . La linealidad sólo puede obtenerse a costa de aproximaciones." Puesto que nuestro interés fundamental no lo constituyen las aplicaciones de la Teoría del Control Óptimo al Análisis de medidas de política económica a corto plazo, que se basarían en un modelo econométrico, sino las aplicaciones al análisis de estrategias de desarrollo a largo plazo, centraremos las consideraciones alrededor de los modelos dinámicos input-output.

Tales modelos son, por su propia naturaleza, lineales, Presentan la indudable ventaja de facilitar datos a un elevado nivel de desagregación y pueden ser utilizados en aplicaciones reales con un menor trabajo previo de estimación de parámetros de lo que requerirían modelos más teóricos.

La forma clásica del modelo dinámico de input-output es :

$$q(t) = Aq(t) + B(q(t+1) - q(t)) + d(t) \quad (1)$$

donde :

$q(t)$ = vector de los niveles de producción en el periodo t .

A = matriz clásica de input-output.

B = matriz de coeficientes de capital.

b_{ij} = stock de capital del tipo i necesario para producir una unidad del bien j .

$d(t)$ = vectores de demanda final.

que puede ser puesta inmediatamente en la forma de ecuaciones de estado de un sistema dinámico :

$$q(t+1) = B^{-1}(I - A + B)q(t) - B^{-1}d(t)$$

donde las variables de estado son los niveles de producción $q(t)$ y los controles los niveles de demanda final.

De acuerdo con esta ecuación, parece que disponemos de un sistema lineal para representar la evolución del sistema económico estudiado. En realidad, las dificultades que la adopción de tal sistema de ecuaciones de estado encierra, son considerables.

Matemáticamente hablando, la hipótesis de la regularidad de B puede ser difícil de mantener. Livesey (17) expone una transformación para superar esta dificultad.

Las objeciones de tipo teórico son mas importantes. Las clasificamos en dos tipos según que provengan de las hipótesis que traduce la identidad contable (1) o de las posibles consecuencias sobre la evolución del sistema de su elevación a la categoría de relación dinámica causal. Ambas se traducen en la necesaria adopción de restricciones desigualdad .

La relación (1) traduce la hipótesis de plena utilización del stock de capital. Solo en estas condiciones es posible concebirla como la ecuación de evolución de un sistema

dinámico, evolución que quedaría así plenamente determinada una vez dados los niveles iniciales de producción y las sucesivas demandas finales $d(t)$, $t=0, 1 \dots t, \dots$ (Morishima(M10))

La relajación de tal hipótesis se expresa mediante la adopción de la desigualdad :

$$K(t) \geq B q(t) \quad (2)$$

donde $K(t)$ es el vector de los stocks de capital.

A su vez, la ecuación (1) pasa a expresarse en la forma:

$$q(t) = A q(t) + K(t+1) - K(t) + d(t)$$

que al sustituirla en (2) origina :

$$K(t) \geq B [I-A]^{-1} [K(t+1) - K(t) + d(t)]$$

$$K(t) + B [I-A]^{-1} K(t) \geq B [I-A]^{-1} [K(t+1) + d(t)]$$

$$[I + B [I-A]^{-1}] K(t) \geq B [I-A]^{-1} [K(t+1) + d(t)]$$

Lo cual, como dice Livesey, es bastante distinto de la relación lineal necesaria para plantear un sistema lineal-cuadrático.

Pero hay más. Un algoritmo numérico de resolución no conoce mas hipótesis económicas que las que aparecen explicitadas en las ecuaciones del modelo que se le somete. Que la teoría económica o el buen sentido común consideren inaceptables determinados valores de las variables de estado y/o control, no implica que estos puedan ser escogidos como óptimos por el algoritmo sino le ha sido expresamente

prohibido el hacerlo.

El modelo anterior no encierra ninguna hipótesis acerca de la irreversibilidad del stock de capital. Y aunque en buena lógica económica un alto horno no puede convertirse de repente en un bien consumible, nada prohíbe al método de cálculo el atribuir a $d(t)$ valores superiores a $q(t)$ a través de la asignación de un valor negativo a $K(t+1)-K(t)$. Tal imposibilidad debe aparecer bajo la forma de una relación del tipo :

$$K(t+1)-K(t) \geq \alpha K(t) \quad \alpha > 0$$

Otro requisito que exigiremos a la solución es que :

$$\begin{aligned} d(t) &\geq 0 \\ q(t) &\geq 0 \quad \forall t \in [0, T] \end{aligned}$$

lo cual tampoco está implícito en la ecuación de estado derivada de (1). Ello ha sido demostrado por Jorgenson (J3), quien prueba que no siempre se generan valores no negativos de los niveles de producción a partir de condiciones iniciales y la aplicación de controles (demanda final), arbitrarios y no negativos. Relaciones de la forma:

$$\begin{aligned} d(t) &\geq 0 \\ q(t) &\geq 0 \quad \forall t \in [0, T] \end{aligned}$$

deben ser introducidos si se desea preservar el realismo de las soluciones.

Tales modificaciones mantienen la linealidad del sistema pero impiden la aplicación directa de los métodos computacionales del modelo lineal cuadrático. El número de restricciones adicionales incorporando es considerable y

disminuye las ventajas inherentes al método de cálculo. .

Para emplear modelos de este tipo como ecuaciones de estado de un sistema económico del que se pretende analizar las estrategias de desarrollo a largo plazo, nos parece necesario modificar las hipótesis que encierra y la linealidad de su expresión analítica.

En el modelo (1), las variables de control son los niveles de demanda final. Así ocurre en los modelos de Smirnov (s5) y Brody (B7). Una vez fijados los valores de estas variables, y dados los niveles de producción $q(t)$ del período actual t , quedan automáticamente determinados los niveles de producción $q(t+1)$ del período siguiente. Lo cual implica que (1) encierra no solamente una explicación del proceso de acumulación de capital. El aceptar relaciones lineales en los procesos de producción y acumulación en un proceso dinámico a largo plazo, parece una hipótesis poco realista.

Para sustituirla y preservar al mismo tiempo el uso de la estructura informativa de un sistema input-output, efectuamos las siguientes modificaciones:

- 1) Desglosar el mecanismo expresado globalmente por la relación (1) en sus tres fases de producción, distribución y acumulación de capital.
- 2) Considerar como variables de control, no los niveles de demanda final, sino las decisiones sectoriales de inversión.

Las relaciones de producción se expresan a través de funciones de producción que permitan incorporar sustitu-

ción entre capital y trabajo y rendimientos variables.

Las necesidades de bienes intermedios para asegurar la factibilidad de los niveles de producción se expresan en las relaciones de distribución. Estas son identidades contables (lineales) que no incorporan efectos de substitución entre bienes, ni entre los consumidos como bienes intermedios ni entre los necesarios para llevar a cabo los niveles de inversión.

El sistema de ecuaciones (1) pasa a ser :

$$q(t) = Aq(t) + BI(t) + d(t) \quad (3)$$

donde $I(t)$ son los niveles sectoriales de inversión.

Dos modificaciones fundamentales separan (3) de (1) :

- Todas las magnitudes relacionadas por (3) están referidas al mismo instante de tiempo. En consecuencia (3) no constituye ninguna ecuación de comportamiento de un sistema dinámico sino una relación de distribución instantánea.
- La matriz B de (3) tiene un significado distinto del atribuido a la matriz B de (1). En (3), los elementos b_{ij} representan la cantidad de bien del tipo i necesario para realizar un nivel de inversión unidad para aumentar el stock de capital del sector j .

Ni (1) ni (3) permiten substitución entre bienes, pero a diferencia de (1), (3) no especifica las cantidades de capital necesarias al proceso productivo, puesto que no se refiere a él. En consecuencia, permite introducir, a

través de la forma que se especifique de las funciones de producción, rendimientos de escala y substitución capital-trabajo.

Finalmente, el tercer sistema de relaciones es el que se establece entre inversión e incremento del stock de capital. Consideramos a este como un proceso sometido a rigideces y a rendimientos variables. El incremento de las dotaciones de capital de los distintos sectores no se verifica automática y linealmente a partir de los niveles de inversión. Las relaciones entre ambas magnitudes son tomadas en cuenta mediante funciones no lineales de la forma :

$$K(t+1) = f(K(t), I(t))$$

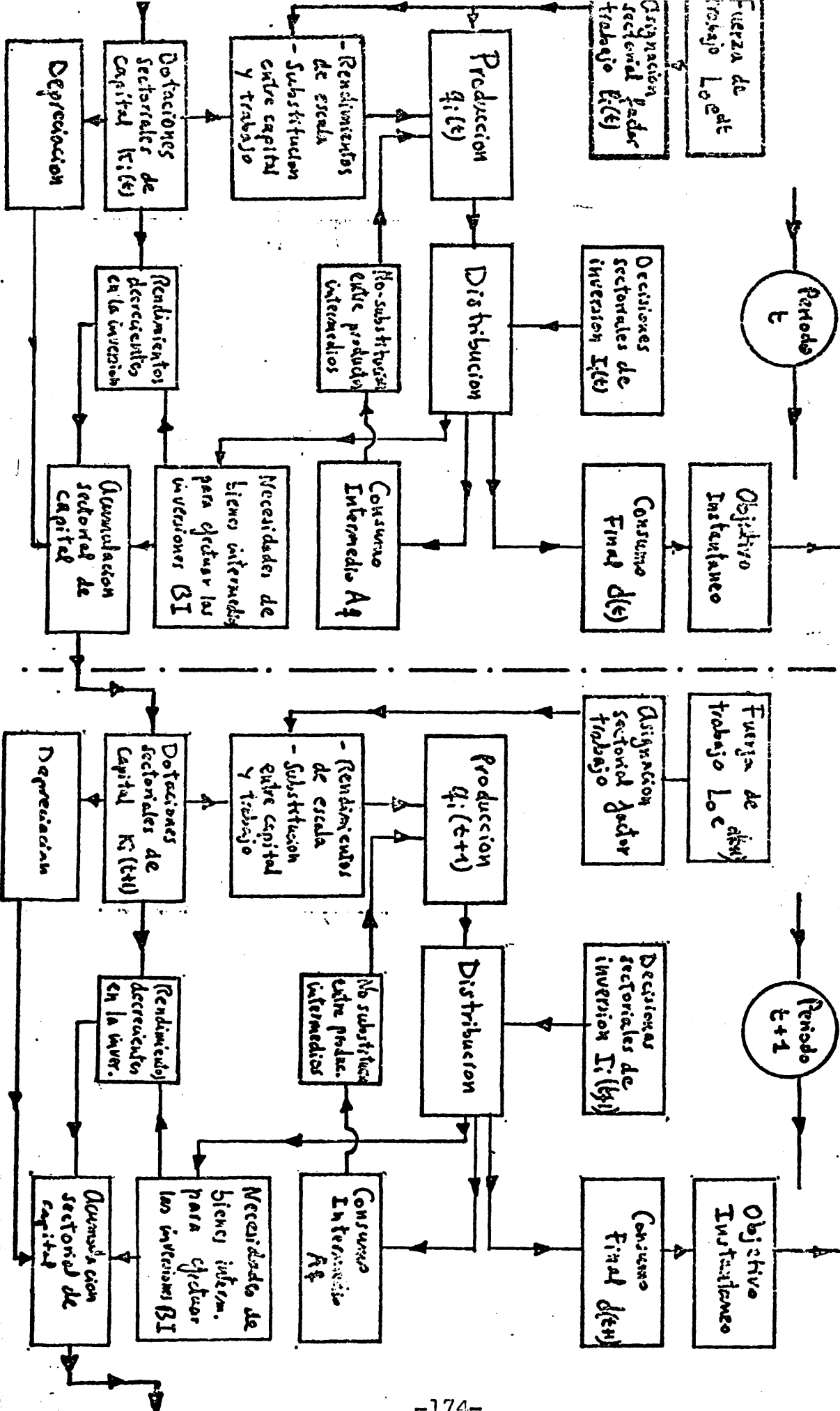
Todas estas consideraciones nos han alejado bastante de nuestro punto de partida. Así haciendolo, se ha cumplido una vez mas la ley general que traduce un mayor realismo de las hipótesis en una mayor complejidad matemática en la formulación del modelo. El esquema del proceso así modelizado es el de la figura F.IV.1.

Limitados como estamos por la eficacia de los métodos y las capacidades de los instrumentos de cálculo, se trata de saber hasta que niveles de complejidad se puede llegar en la formulación manteniendo la posibilidad de resolver el modelo construido.

Teniendo esto presente, el punto IV. 3 desarrolla más extensamente las anteriores consideraciones, en relación con los algoritmos del capítulo III.

Antes de entrar en ello, analicemos las características del segundo componente de un modelo lineal-cuadrático.

FUNCION OBJETIVO GLOBAL



IV. 2. 3 CARACTERISTICAS DE LAS FUNCIONES OBJETIVO CUADRÁTICAS.

Como hemos recordado, un objetivo cuadrático, asociado con ecuaciones de estado lineales, permite la aplicación del principio de certeza-equivalencia y posee una conocida forma de resolución.

Además, la función objetivo cuadrática, cuya forma mas simple es:

$$f(x) = a_0 + a_1 x + a_2 x^2, \quad a_2 < 0$$

presenta, por si misma, interesantes propiedades:

Permite prescindir de la constancia de las magnitudes marginales, inherente a las relaciones lineales, y contiene una penalización implícita de las desviaciones de x con respecto a su valor óptimo $-\frac{a_1}{a_2}$ como lo muestra la forma equivalente:

$$f(x) = \left(\frac{a_0}{2} - \frac{a_1^2}{2a_2} \right) + \frac{1}{2} a_2 \left[x - \left(-\frac{a_1}{a_2} \right) \right]^2$$

La penalización de las desviaciones permite, en primera aproximación, prescindir de restricciones sobre las variables.

Cuando hay mas de una variable hay que ponderar relativamente las desviaciones de cada una de ellas con respecto a su respectiva trayectoria objetivo mediante los elementos de las matrices A, B :

$$\sum_{K=0}^N \left[\|x(k) - \hat{x}(k)\|_A^2 + \|u(k) - \hat{u}(k)\|_B^2 \right]$$

Indisolublemente asociados con estas ventajas, aparecen una serie de inconvenientes de mucho mayor peso en las aplicaciones económicas que en el análisis de problemas físicos. Podemos clasificarlos en dos grandes grupos según que su componente principal sea de cálculo numérico o de definición política.

La aplicación del principio de certeza-equivalencia es una ventaja relativa cuando los elementos aleatorios no se limitan a términos aditivos en las ecuaciones de estado o de medida. Parece evidente que en la descripción de sistemas económicos los coeficientes de las matrices de ambos sistemas de ecuaciones son tan susceptibles de contener incertidumbre como el mismo proceso. Sin embargo, una de las condiciones de aplicación del principio de certeza-equivalencia es la de que tales elementos, aún sin ser constantes, tengan una evolución determinista. Para un análisis de errores, ver (Terceiro (T4)).

La penalización cuadrática es por definición simétrica y no distingue entre desviaciones por defecto y por exceso de las variables con respecto a la trayectoria objetivo. Ello implica que nuestro sistema de valoración es tan sensible a un deficit de la balanza de pagos como a un superavit ; o que si el nivel deseado de desempleo es el 3% penalizamos igualmente una política que produzca un 4% como una que nos permita alcanzar el 2%. La valoración de las desviaciones , siendo esencialmente asimétrica en economía, dos tipos de razonamientos han sido utilizados para justificar

el uso de un modelo intrínsecamente simétrico.

- La fatalista (Pindyck (P15)). Las cosas suelen ir mal: los niveles de desempleo y de inflación, serán superiores a los deseados y los de renta y reservas inferiores. La vertiente ilógica del sistema de valoración no tendrá, pues, ocasión de actuar.
- La positiva (Friedman (F8)). Adoptando una función cuadrática por intervalos con coeficientes de signo y valor variable en función de la zona en la que se sitúe las desviaciones. El inconveniente se salva a costa de una mayor complejidad en los algoritmos de solución.

La mas seria dificultad proviene sin duda alguna de la doble elección que obliga a realizar a los responsables políticos. Por una parte, deben describir explícitamente las trayectorias deseadas del sistema económico $(\hat{x}(k), \hat{u}(k))$ y por otra, deben definir su valoración relativa de las desviaciones de las variables alrededor de su "trayectoria objetivo" (matrices A y B).

Una definición coherente de su posición a estos respectos es más de lo que un responsable político sea capaz de hacer. Más, incluso, de lo que desee hacer.

En la práctica, tal tipo de información es imposible de obtener y se substituye por una serie de análisis con coeficientes alternativos representantes de distintos sistemas de preferencia, analizando así las consecuencias de los mismos. Merecen destacarse las posibilidades, tan variadas como interesantes , que este procedimiento ofrece para tratar

de ilustrar los compromisos("trade offs") inherentes a todo plan. Se puede, por ejemplo, analizar el coste que representaría el seguir exactamente la trayectoria objetivo de una de las variables de estado. Ello se instrumentaría mediante valores muy elevados en los coeficientes de la matriz A correspondientes a dicha variable. Tal análisis implica un elevado número de "runs" del modelo con su coste correspondiente.

Sin embargo, el elemento de más difícil definición es, sin duda, la propia trayectoria objetivo con respecto a la cual se trata de minimizar las desviaciones de la trayectoria real óptima. Puesto que esta es un dato previo al proceso de optimización, este puede convertirse en una esteril pugna para hacer seguir al sistema una trayectoria de desarrollo que no sea en absoluto aquella que debiera seguir en el óptimo.

Encontramos aquí el razonamiento con el que empezamos este capítulo : la aplicación de los métodos de control óptimo al análisis de las estrategias de desarrollo a largo plazo tiene por objeto el determinar la trayectoria temporal óptima en el sentido definido por una cierta funcional. El modelo matemático empleado difícilmente puede ser el lineal cuadrático puesto que este precisa conocer ya la trayectoria que se debe tratar de seguir.

Una vez determinada la trayectoria óptima mediante un método general de control óptimo, queda abierta la aplicación del modelo lineal cuadrático al análisis de su óptima realización. Las características del modelo lineal-cuadrático le hacen especialmente adaptado al análisis, a corto plazo, de los elementos aleatorios que inciden sobre el.

sistema. Por ello, el modelo lineal-cuadrático es el instrumento ideal para el control óptimo a corto plazo de medidas de política monetaria y fiscal sobre modelos econométricos lineales.

Finalmente, no todas las funciones objetivo susceptibles de ser escogidas (o de ser impuestas) en el análisis de las estrategias de desarrollo a largo plazo se expresan mediante una funcional (cuadrática o no) de estados y controles. Una amplia clase de objetivos son los llamados de mínimo tiempo. Con ellos se trata de determinar la forma más rápida de alcanzar un determinado estado a partir de la situación actual del sistema. No se efectúa evaluación alguna de lo que ocurra durante la trayectoria, siempre que esta se mantenga dentro de los límites fijados por un conjunto de restricciones instantáneas. Este tipo de objetivos parecen especialmente interesantes en su aplicación a la planificación del proceso de "despegue" de economías subdesarrolladas. Sin entrar en un análisis más extenso del tema, citemos simplemente el extenso tratamiento de que está siendo objeto recientemente en la literatura soviética(ДРУКАЛОВ(D15), ДУБОВСКИЙ(D14, D13)).

Evidentemente, un problema de mínimo tiempo no tiene sentido sin la inclusión de un conjunto de restricciones sobre el estado final. Pero la definición de un dominio de estados finales aceptables y de un sistema de valoración asociado a este, es perfectamente complementario de objetivos expresados mediante una funcional de estado y controles.

Ambos elementos son tomados en consideración en los

métodos de control óptimo en general y en los algoritmos del capítulo III en particular. Puesto que el horizonte del plan es necesariamente finito en todo problema real, es preciso complementar la formulación del objetivo mediante la imposición de restricciones sobre el estado final del sistema, que expresen la preocupación (elemento aún de definición política) por las trayectorias potenciales del sistema después de T.

Sin pretender recoger todas las conclusiones a las que hemos ido llegando, podemos resumir diciendo que :

La formulación lineal-cuadrática en modelos multi-sectoriales no excluye, aunque sí simplifique, la aplicación de métodos numéricos.

La ruptura de la simetría en la valoración introduce una complicación adicional.

Las dificultades de definición de tipo político que están en la base del modelo obligan a un análisis de multitud de alternativas.

Este triple serie de circunstancias hace pensar que una formulación mas general con inclusión de restricciones aliaría a su mayor realismo una, al menos comparable, capacidad operativa.

Ciertamente la inclusión de tales restricciones y el cambio de forma en la función objetivo obliga a adoptar otros procedimientos de cálculo. Pero permite a la vez una definición mas realista y coherente de las prioridades políticas, sin la cual los resultados numéricos carecen de

sentido.

La hipótesis así enunciada constituye una de las actuales controversias en planificación económica. Solo el desarrollo de las, por el momento escasas, experiencias computacionales, determinará su validez.

Ello no va sin reconocer el interés del modelo lineal-cuadrático en sus aplicaciones económicas:

Descomponiendo el proceso de control del sistema económico en dos fases :

- una determinista con una formulación general que determine la trayectoria óptima de desarrollo a largo plazo.
- otra, estocástica, de regulación de la trayectoria real alrededor de esta trayectoria óptima nominal,

El sistema lineal-cuadrático, por :

- su relación con los sistemas de regulación.
- presuponer la existencia de una senda que se desea siga la economía.
- la aplicación del teorema de certeza-equivalencia para tratar determinados tipos de elementos aleatorios.

parece el instrumento adecuado para ser aplicado en esta segunda fase, es decir, en el análisis de medidas instrumentales de política coyuntural, estabilización y, en general, política económica a corto plazo.

IV. 3. 1 PLANTEAMIENTO DEL MODELO.

Pretendemos aquí aplicar los métodos y conceptos de anteriores capítulos a la elaboración de un modelo que analice cuantitativa y desagregadamente las estrategias de desarrollo a largo plazo de una economía en desarrollo. Dicha economía se enfrenta con una necesidad de financiación exterior, problemas de pleno empleo y bajos niveles de capitalización y consumo. Sus objetivos a largo plazo, decididos políticamente y bajo el control de un hipotético centro único de decisión, se expresan a través de su valoración de tal problemática mediante una única funcional. El análisis de sus estrategias de actuación requiere la consideración de los distintos "trade-offs" entre las trayectorias de consumo e inversión, el reparto de ambas categorías entre los distintos sectores de actividad y las alternativas entre capital y trabajo en los procesos productivos.

La toma en consideración de relaciones no lineales permite un análisis mas realista de tales alternativas. El carácter dinámico del problema obliga a considerar condiciones terminales, al tiempo que la optimalidad de las trayectorias se juzga globalmente sobre todo el periodo considerado. La información dual de las variables adjuntas permitirá analizar los equilibrios temporales realizados.

El modelo es del tipo propuesto por Kerdrick y Taylor, (K7), (K6). Como el mismo Kerdrick menciona, la novedad y escasa experiencia computacional en modelos de planificación dinámica no lineales hace que el interés actual de tal tipo de modelos, se centre mas en el análisis de sus

posibilidades que en su aplicación a una situación concreta.

Las principales diferencias entre nuestro planteamiento y el original de Kendrick son:

- a) la adopción de una funcional objetivo mas completa.
- b) el abandono de la hipótesis de pleno empleo.
- c) toma en consideración de restricciones instantáneas sobre las variables, introduciendo así condicionamientos políticos que no se expresan directamente en la funcional objetivo y garantizando el realismo de las soluciones obtenidas.
- d) cálculo y análisis de la información dual.

Tales modificaciones son sólo posibles gracias a la aplicación de los métodos computacionales del capítulo III. El propio Kendrick hace referencia a la necesidad de incorporar hipótesis poco realistas para obtener un modelo tratable numericamente. La posibilidad de tomar en cuenta explícitamente restricciones desigualdad (grave fuente de complicaciones de cálculo, inabordables en el método utilizado por Kendrick) es la que hace posible abandonar hipótesis tan ideales como la de pleno empleo y las posibilidades de endeudamiento externo instantáneo no acotadas.

En la exposición que sigue de la estructura del modelo, analizamos sucesivamente sus elementos constituyentes, indicando en detalle las modificaciones que hemos introducido en el modelo original.

1) FUNCIONAL OBJETIVO .

Tres son los elementos que tomamos en consideración en la construcción de la funcional objetivo:

- los niveles instantáneos actualizados de consumo de los bienes producidos por cada uno de los sectores.
- los niveles de desempleo soportados.
- la capacidad productiva con que la economía se encuentra dotada al final del periodo considerado.

Tales elementos son combinados de la siguiente forma:

$$J = \sum_{t=1}^T \left[(1+\rho)^{-t} \sum_{i=1}^N a_i [C_i(t)]^{b_i} - \alpha \left[\ell(t) - \sum_{i=1}^N \ell_i(t) \right] \right] + \sum_{i=1}^N \nu \tau K_i (K_i(T) - K_i^*)$$

En donde:

T : número de periodos de tiempo considerados.

N : número de sectores considerados.

ρ : tipo de actualización social.

$C_i(t)$: nivel de consumo del bien producido en el sector i en el periodo t .

a_i : parámetro $a_i > 0$

b_i : parámetro $0 \leq b_i \leq 1$ $i=1, N$

α : parámetro de ponderación de los efectos del desempleo en el objetivo. (elección política).

$L(t)$: fuerza de trabajo total disponible en el periodo t .

$l_i(t)$: nivel de empleo en el sector i en el instante t .

VTK_i : peso otorgado a la disponibilidad de una unidad suplementaria de capital en el sector i en el instante final del periodo del plan, por encima del nivel K_i^*

K_i^* : nivel mínimo de capital que se exige haber dotado al sector i al final del plan.

Como se observa, tal funcional objetivo debe venir acompañada de la especificación del criterio terminal K_i^* , es decir:

2) CONDICIONES TERMINALES.

$$K_i(T) \geq K_i^*$$

La forma misma del objetivo es un acto de definición política. Obsérvese que en el se incluyen tres de las formas en que se puede expresar el objetivo en un problema standard de control óptimo : definición de un "blanco" terminal, de una valoración del estado final del sistema y de una evaluación de su trayectoria.

En el modelo propuesto por Kendrick, solamente se tiene en cuenta como objetivo la suma actualizada del consumo. Las condiciones terminales aparecen planteando el problema como uno de estado final fijado:

$$K_i(\tau) = K_i^* \quad (i=1, N)$$

En la formulación adoptada, los parámetros $\rho, \alpha, \sum VTK_i$ ($i=1, N$), K^* son elementos definidos políticamente. Sin ser los únicos con tal origen, aparece evidente que sus valores relativos son de importancia fundamental. Revelan las preferencias entre los distintos objetivos englobados en \sum y definen el sentido del adjetivo "óptimo" asignado a la trayectoria seleccionada por el proceso de cálculo.

Las expresiones $\sum VTK_i (K_i(\tau) - K_i^*)$ y $K(\tau) \geq K^*$ expresan la preocupación por las trayectorias potenciales del sistema después del fin del periodo del plan. Algunos modelos propuestos se limitan a considerar exclusivamente objetivos de este tipo. Preferimos completarlo con una evaluación de la trayectoria en su desarrollo temporal puesto que la más altruista de las sociedades no cifra únicamente su interés en lo que ocurra en ó después de una fecha más o menos lejana. Esta preocupación aparecerá de nuevo cuando explicitemos las condiciones sobre la trayectoria.

Puesto que es sabido que en modelos no lineales con utilidad marginal decreciente y evolución de la población exógenamente definida, la maximización del consumo global no produce los mismos resultados que la maximización del consumo per cápita (K_8), importa señalar que las magnitudes

$C_i(t)$ no representan consumo per cápita.

Con los valores de los parámetros a_i, b_i la función objetivo presenta:

- utilidad marginal positiva pero decreciente para cada bien i en cada periodo t :

$$\frac{\partial J}{\partial C_i(t)} = (1+r)^{1-t} a_i b_i \frac{1}{[C_i(t)]^{1-b_i}}$$

- separabilidad entre los efectos de los distintos bienes y distintos periodos :

$$\frac{\partial J}{\partial C_i(t) \partial C_j(t)} = 0$$

$$\begin{aligned} \forall i \neq j \\ i=1, N \\ j=1, N \end{aligned}$$

$$\frac{\partial J}{\partial C_i(t) \partial C_i(t+\theta)} = 0$$

$$\forall \theta \neq 0 \quad \forall t \in [1, T]$$

La consideración explícita de utilidades marginales decrecientes es de extrema importancia en un análisis a largo plazo y su defecto puede introducir importantes distorsiones en el realismo de los resultados obtenidos. Esta es una de las mas importantes ventajas que pueden obtenerse del abandono de la linealidad. Evidentemente, la hipótesis de separabilidad no es deseable, pero sus efectos son menos nocivos, sobre todo a los niveles de agregación usada, y la posibilidad de abandonarla precisa solamente reformular J . Ello no introduciría dificultades de cálculo suplementario con los métodos empleados.

Como exponen Kendrick y Taylor (K6) y Mehra (M6), una ventaja adicional de tal tipo de función es la de permitir derivar procedimientos para la estimación de los parámetros a_i, b_i .

La descripción del estado inicial pasa por y conduce a la elección de las variables de estado del sistema.

Los elementos que escogemos para caracterizar la situación inicial son:

- el capital con el que se encuentran dotados cada uno de los sectores : $K_i(1) \quad (i=1,N)$
- la situación de endeudamiento externo del país en el instante inicial del plan : $\chi(1)$
- la fuerza de trabajo disponible en el instante inicial $l(1)$.

Aparece aquí una situación que es preciso analizar en detalle : A lo largo del modelo, la fuerza de trabajo asignada a los distintos sectores será considerada como una variable de control puesto que tal asignación es, dentro de ciertos límites, una de las posibles alternativas de acción.

Sin embargo, no es menos evidente que tal distribución sectorial de la fuerza de trabajo es, en el instante inicial, un dato impuesto por la realidad. Por lo tanto, en la descripción inicial del sistema es preciso incluir las magnitudes $l_i(1), i=1,N$.

Puesto que nos proponemos tratar explícitamente las situaciones de desempleo, aceptamos que $\sum_{i=1,N} l_i(1) \leq l(1)$, es decir, la distribución inicial de la fuerza de trabajo entre los sectores no agota a esta. Las magnitudes $l_i(1)$ representan los niveles sectoriales reales de empleo. Ello nos lleva directamente al espinoso problema del tratamiento analítico del paro encubierto y del subempleo. Evidentemente, la variable desempleo (DSMP (t)) no existe físicamente en una antesala de espera del trabajo, sino que coexiste en los distintos sectores con la parte de la fuerza de

trabajo realmente empleada en ellos. Parece sin embargo analíticamente superfluo desagregar sectorialmente tal magnitud. En consecuencia, las magnitudes $l_i(t)$ representarán los niveles de trabajo realmente empleados en los distintos sectores de acuerdo con los límites impuestos por las disponibilidades de capital a través de las correspondientes funciones de producción. La variable $DSMP(t) = l(t) - \sum_{i=1, N} l_i(t)$ juega el papel analítico de variable de holgura en la restricción desigualdad que substituye a la hipótesis neoclásica de pleno empleo. Su valor es de importantes consecuencias en la valoración de las trayectorias a través de la funcional J .

La descripción de la situación inicial debe pues completarse con las magnitudes $l_i(1), (i=1, N)$ de las que se deduce inmediatamente el nivel inicial de desempleo $DSMP(1)$.

Quedan así definidos el objetivo y las condiciones iniciales y finales del sistema. El problema es guiar óptimamente su evolución entre ambas situaciones. Para ello es preciso describir los elementos de su dinámica y las posibilidades de actuación sobre la misma de las que se dispone. En otras palabras, es preciso definir las ecuaciones de estado y las variables de control.

Las variables de estado ya han sido escogidas y, de acuerdo con el punto anterior, son:

$\delta(t)$: deuda exterior en el periodo t .

$K_i(t)$: capital del sector i en el instante t .
($i=1, N$)

Para tener en cuenta los efectos de rendimientos variables, substitución entre factores, no linealidades en los procesos de acumulación, etc..., las ecuaciones de estado son bastante complejas. Su significado económico queda mas claramente expuesto si las describimos descompuestas en las fases que representan los procesos de producción, distribución, acumulación de capital y la correspondiente modificación de la situación de la balanza exterior y niveles de empleo.

4) PROCESO PRODUCTIVO.

Tal como ha sido caracterizado el sistema, las cantidades de capital y trabajo empleadas en cada sector son magnitudes agregadas, lo que implica que estamos considerando una única clase de capital y de trabajo por sector.

A fin de considerar un mayor nivel de desagregación en los inputs intermedios, coherente con la división en sectores, las funciones de producción empleadas no consideran explícitamente sino los inputs básicos de capital y trabajo. La necesaria utilización en el proceso productivo de inputs intermedios se toma en cuenta a través de las ecuaciones de distribución de la producción.

Las funciones de producción empleadas deben traducir dos importantes fenómenos:

- substitución entre capital y trabajo.
- rendimientos de escala variables.

El poder tener en cuenta la substitución de capital por trabajo es absolutamente necesario en modelos aplicados a economías en desarrollo con fuertes niveles de paro. Como Kendrick, adoptamos una función de producción con elasticidad de substitución constante, pero distinta de 1 y variable de sector a sector. Pero completamos tal formula-

ción de forma que se pueden considerar simultáneamente rendimientos de escala variables y que se tenga en cuenta los límites físicos a la substitución entre factores.

Las funciones consideradas son de la forma :

$$q_i(t) = \eta_i (1 + \gamma_i)^{t-1} \left[\beta_i [k_i(t)]^{-\rho_i} + (1 - \beta_i) [l_i(t)]^{-\rho_i} \right]^{-\frac{\theta_i}{\rho_i}}$$

$(i = 1, N)$

en donde las variables representan:

$q_i(t)$: producción del sector i en el periodo (t)

$\eta_i(t)$: parámetro característico de cada sector
(llamado parámetro de eficiencia).

γ_i : factor de progreso técnico.

β_i : parámetro de distribución.

$k_i(t)$: cantidad de capital empleado en el periodo
 t en el proceso productivo del sector i .

$l_i(t)$: nivel de empleo en el sector i en el
periodo t .

$$\rho_i = \frac{1}{\sigma_i} - 1$$

σ_i : elasticidad de substitución entre capital
y trabajo en el sector i .

θ_i : parámetro representativo de los rendi-
mientos de escala.

Es preciso hacer las siguientes consideraciones acerca de tal función:

- el progreso técnico es considerado de forma muy simple. Se le considera neutro con respecto a capital y trabajo y función solamente del paso del tiempo.
- las posibilidades de substitución entre factores dependen del sector al que se refiera la función.
- suponemos que no habrá nunca un exceso de capacidad productiva disponible en el factor capital. Por ello utilizamos el mismo símbolo $(K_i(t))$ para representar el capital de cada sector y la cantidad efectivamente empleada del mismo. Tal hipótesis parece concomitante con el tipo de situación que estamos analizando.
- cuando se consideran rendimientos de escala crecientes $(\theta > 1)$, la forma adoptada por la función de producción puede hacer que el dominio de soluciones no sea convexo. Ello no debe ser obstáculo a la introducción de tal hipótesis puesto que la convexidad es en sí destruida por la existencia de restricciones igualdad no lineales y no es un requisito fundamental para la aplicación de algunos de los algoritmos de cálculo propuestos.
- tal como esta formulada, la función de producción permite, en cada sector, la substitución en el proceso productivo, de todo el capital por trabajo y vice-versa.

Ello es evidentemente ideal. De mantenerse tal hipótesis sería superfluo preocuparse por el problema de pleno empleo, puesto que en el análisis numérico del modelo, toda la fuerza de trabajo sería inmediatamente ocupada aún con cantidades nulas de capital.

Si bien es preciso recordar que los modelos de desarrollo que, como el chino, mas éxito han alcanzado, incorporan al hombre a la producción con dotaciones mínimas de capital,

parece evidente que:

- no es realista suponer que grandes trasvases de mano de obra pueden efectuarse de sector a sector en un corto periodo de tiempo.
- la fuerza de trabajo que es posible emplear en determinados sectores de actividad, por ejemplo la industria pesada, está acotada por el capital disponible.

Sin suponer que la expresión analítica de tal cota deba pasar por el origen, es preciso completar la función de producción propuesta mediante un conjunto de restricciones de la forma:

$$l_i(t) \leq f_i(K_i(t)) \quad (i=1, N) \quad t \in [1, T]$$

cuya expresión particular dependerá del sector considerado.

Tales restricciones son un complemento imprescindible de las funciones de producción definidas. Sin ellas, la restricción $\sum_{i=1}^N l_i(t) \leq l(t)$ quedaría sistemáticamente saturada, lo que equivaldría a aceptar implícitamente la hipótesis de pleno empleo y la imposibilidad de analizar las relaciones entre empleo y acumulación de capital a través de las distintas alternativas de producción.

En efecto, las funciones de producción empleadas tienen la indeseable propiedad de generar productividades marginales del trabajo decrecientes pero nunca nulas:

$$\left(\frac{\partial q(t)}{\partial l_i(t)} \right)_{K_i(t)=K^*} = \frac{Q_i}{f_i} \eta_i (1+\nu_i)^{t-1} M_1 \cdot M_2 \neq 0 \quad \forall l_i, K^* > 0$$

$$M_1 = \left[\beta_i [K_i(t)]^{-\rho_i} + (1-\beta_i) [l_i(t)]^{-\rho_i} \right]^{-\frac{Q_i}{\rho_i} - 1}$$

$$M_2 = f_i(1-\beta_i) [l_i(t)]^{-\beta_i-1}$$

Es decir, las cantidades del factor trabajo que pueden ser incorporadas eficazmente al proceso productivo con una misma dotación de capital, son ilimitadas e independientes del valor de estas dotaciones. Gracias a esta ficción matemática, no solamente se consigue hacer desaparecer el problema del desempleo sino también la posibilidad analítica de tomarlo en consideración.

El que se trata de una ficción matemática aparece evidente cuando se pasen de las expresiones analíticas a los valores numéricos. Como muestran las figuras F.IV.2,3, las productividades marginales del trabajo tienden a cero rápidamente y a partir de un cierto valor de $l_i(t)$ (que depende del valor de $K_i^*(t)$ para el que han sido calculadas) pueden considerarse nulas.

Las productividades marginales del factor trabajo representados en dichas figuras han sido calculadas para dos conjuntos distintos de dotaciones de capital en 4 sectores de actividad. Los parámetros de las correspondientes funciones de producción son los dados en el modelo de Kendrick y Taylor (Figura no.IV.2,3) (K6).

Aceptando que para cada valor de $K_i^*(t)$, la productividad marginal del trabajo puede considerarse nula a partir de un cierto $l_i^*(t)$, del anterior razonamiento puede deducirse inmediatamente una forma de las restricciones

$$l_i(t) \leq f(K_i(t)) \text{ que sea coherente con los demás datos del modelo.}$$

$$f(K_i(t))$$

En el instante inicial se suponen conocidas las combina-

PRODUCTIVIDADES MARGINALES DEL FACTOR TRABAJO
CALCULADAS PARA LAS SIGUIENTES DOTACIONES DE CAPITAL
.....

	1.0000.+01	1.5300.+01	8.5000.+00	7.5000.+00
TRABAJO	AGRICULTURA	IND.PESADA	IND.LIGERA	SERVICIOS
1	1.4976.+00	3.6105.+00	4.0488.+00	6.5031.-01
2	1.1076.+00	3.0878.+00	3.5393.+00	6.3723.-01
3	9.3553.-01	2.8051.+00	3.2584.+00	6.2655.-01
4	8.3280.-01	2.6150.+00	3.0670.+00	6.1722.-01
5	7.6245.-01	2.4734.+00	2.9231.+00	6.0883.-01
6	7.1032.-01	2.3615.+00	2.8085.+00	6.0112.-01
7	6.6962.-01	2.2696.+00	2.7138.+00	5.9396.-01
8	6.3628.-01	2.1919.+00	2.6334.+00	5.8724.-01
9	6.0928.-01	2.1250.+00	2.5636.+00	5.8090.-01
10	5.8601.-01	2.0662.+00	2.5022.+00	5.7488.-01
11	5.6591.-01	2.0141.+00	2.4474.+00	5.6914.-01
12	5.4831.-01	1.9673.+00	2.3961.+00	5.6365.-01
13	5.3272.-01	1.9249.+00	2.3532.+00	5.5838.-01
14	5.1878.-01	1.8862.+00	2.3122.+00	5.5331.-01
15	5.0622.-01	1.8507.+00	2.2744.+00	5.4842.-01
16	4.9432.-01	1.8178.+00	2.2394.+00	5.4370.-01
17	4.8441.-01	1.7874.+00	2.2068.+00	5.3913.-01
18	4.7484.-01	1.7590.+00	2.1764.+00	5.3470.-01
19	4.6602.-01	1.7325.+00	2.1479.+00	5.3040.-01
20	4.5784.-01	1.7076.+00	2.1211.+00	5.2623.-01
21	4.5023.-01	1.6841.+00	2.0957.+00	5.2217.-01
22	4.4312.-01	1.6619.+00	2.0718.+00	5.1821.-01
23	4.3647.-01	1.6409.+00	2.0491.+00	5.1436.-01
24	4.3021.-01	1.6210.+00	2.0274.+00	5.1060.-01
25	4.2433.-01	1.6021.+00	2.0069.+00	5.0693.-01
26	4.1877.-01	1.5840.+00	1.9872.+00	5.0334.-01
27	4.1351.-01	1.5668.+00	1.9684.+00	4.9984.-01
28	4.0852.-01	1.5503.+00	1.9504.+00	4.9641.-01
29	4.0370.-01	1.5346.+00	1.9332.+00	4.9305.-01
30	3.9927.-01	1.5194.+00	1.9166.+00	4.8977.-01
31	3.9496.-01	1.5049.+00	1.9007.+00	4.8655.-01
32	3.9086.-01	1.4909.+00	1.8853.+00	4.8339.-01
33	3.8693.-01	1.4775.+00	1.8705.+00	4.8030.-01
34	3.8317.-01	1.4645.+00	1.8562.+00	4.7726.-01
35	3.7956.-01	1.4520.+00	1.8424.+00	4.7428.-01
36	3.7610.-01	1.4399.+00	1.8291.+00	4.7136.-01
37	3.7277.-01	1.4282.+00	1.8162.+00	4.6848.-01
38	3.6957.-01	1.4169.+00	1.8037.+00	4.6566.-01
39	3.6649.-01	1.4060.+00	1.7915.+00	4.6288.-01
40	3.6351.-01	1.3954.+00	1.7797.+00	4.6015.-01
41	3.6064.-01	1.3851.+00	1.7683.+00	4.5747.-01
42	3.5787.-01	1.3751.+00	1.7572.+00	4.5482.-01
43	3.5519.-01	1.3654.+00	1.7464.+00	4.5222.-01
44	3.5260.-01	1.3559.+00	1.7359.+00	4.4967.-01
45	3.5009.-01	1.3467.+00	1.7257.+00	4.4715.-01
46	3.4766.-01	1.3378.+00	1.7157.+00	4.4466.-01
47	3.4530.-01	1.3291.+00	1.7060.+00	4.4222.-01
48	3.4301.-01	1.3206.+00	1.6965.+00	4.3981.-01
49	3.4079.-01	1.3124.+00	1.6873.+00	4.3744.-01
50	3.3863.-01	1.3043.+00	1.6782.+00	4.3510.-01
51	3.3653.-01	1.2965.+00	1.6694.+00	4.3279.-01
52	3.3449.-01	1.2888.+00	1.6608.+00	4.3052.-01
53	3.3251.-01	1.2813.+00	1.6524.+00	4.2827.-01
54	3.3058.-01	1.2740.+00	1.6442.+00	4.2606.-01
55	3.2869.-01	1.2668.+00	1.6361.+00	4.2387.-01
56	3.2686.-01	1.2598.+00	1.6282.+00	4.2172.-01
57	3.2507.-01	1.2529.+00	1.6205.+00	4.1959.-01
58	3.2332.-01	1.2462.+00	1.6130.+00	4.1749.-01
59	3.2162.-01	1.2397.+00	1.6056.+00	4.1542.-01
60	3.1996.-01	1.2332.+00	1.5983.+00	4.1337.-01

Figura IV.2.

PRODUCTIVIDADES MARGINALES DEL FACTOR TRABAJO
CALCULADAS PARA LAS SIGUIENTES DOTACIONES DE CAPITAL
.....

	2.0200.+00	2.1300.+00	1.2600.+00	1.2700.+00
TRABAJO	AGRICULTURA	IND. PESADA	IND. LIGERA	SERVICIOS
1	7.6543.-01	2.2544.+00	2.7364.+00	6.0182.-01
2	5.8816.-01	1.8728.+00	2.3326.+00	5.6467.-01
3	5.0801.-01	1.6718.+00	2.1149.+00	5.3594.-01
4	4.5941.-01	1.5387.+00	1.9687.+00	5.1201.-01
5	4.2575.-01	1.4409.+00	1.8600.+00	4.9131.-01
6	4.0058.-01	1.3644.+00	1.7743.+00	4.7301.-01
7	3.8079.-01	1.3021.+00	1.7027.+00	4.5657.-01
8	3.6468.-01	1.2498.+00	1.6444.+00	4.4163.-01
9	3.5120.-01	1.2050.+00	1.5932.+00	4.2794.-01
10	3.3959.-01	1.1659.+00	1.5483.+00	4.1530.-01
11	3.2971.-01	1.1314.+00	1.5065.+00	4.0358.-01
12	3.2094.-01	1.1006.+00	1.4727.+00	3.9264.-01
13	3.1314.-01	1.0728.+00	1.4404.+00	3.8239.-01
14	3.0615.-01	1.0475.+00	1.4109.+00	3.7277.-01
15	2.9983.-01	1.0244.+00	1.3838.+00	3.6369.-01
16	2.9407.-01	1.0032.+00	1.3588.+00	3.5511.-01
17	2.8880.-01	9.8348.-01	1.3356.+00	3.4690.-01
18	2.8395.-01	9.6520.-01	1.3140.+00	3.3926.-01
19	2.7946.-01	9.4315.-01	1.2937.+00	3.3191.-01
20	2.7529.-01	9.3219.-01	1.2748.+00	3.2491.-01
21	2.7141.-01	9.1719.-01	1.2569.+00	3.1822.-01
22	2.6777.-01	9.0307.-01	1.2401.+00	3.1183.-01
23	2.6436.-01	8.8974.-01	1.2251.+00	3.0570.-01
24	2.6115.-01	8.7711.-01	1.2090.+00	2.9983.-01
25	2.5812.-01	8.6513.-01	1.1946.+00	2.9419.-01
26	2.5525.-01	8.5373.-01	1.1809.+00	2.8877.-01
27	2.5254.-01	8.4288.-01	1.1678.+00	2.8356.-01
28	2.4996.-01	8.3253.-01	1.1553.+00	2.7854.-01
29	2.4750.-01	8.2263.-01	1.1433.+00	2.7370.-01
30	2.4516.-01	8.1315.-01	1.1318.+00	2.6903.-01
31	2.4293.-01	8.0406.-01	1.1208.+00	2.6452.-01
32	2.4080.-01	7.9534.-01	1.1102.+00	2.6016.-01
33	2.3875.-01	7.8696.-01	1.0999.+00	2.5595.-01
34	2.3679.-01	7.7889.-01	1.0901.+00	2.5187.-01
35	2.3491.-01	7.7112.-01	1.0806.+00	2.4792.-01
36	2.3310.-01	7.6363.-01	1.0714.+00	2.4410.-01
37	2.3135.-01	7.5639.-01	1.0626.+00	2.4039.-01
38	2.2968.-01	7.4940.-01	1.0540.+00	2.3679.-01
39	2.2807.-01	7.4264.-01	1.0457.+00	2.3329.-01
40	2.2651.-01	7.3610.-01	1.0377.+00	2.2990.-01
41	2.2500.-01	7.2976.-01	1.0299.+00	2.2660.-01
42	2.2354.-01	7.2362.-01	1.0223.+00	2.2340.-01
43	2.2213.-01	7.1766.-01	1.0149.+00	2.2028.-01
44	2.2077.-01	7.1188.-01	1.0078.+00	2.1725.-01
45	2.1944.-01	7.0626.-01	1.0009.+00	2.1429.-01
46	2.1816.-01	7.0079.-01	9.9409.-01	2.1142.-01
47	2.1691.-01	6.9548.-01	9.8750.-01	2.0861.-01
48	2.1570.-01	6.9031.-01	9.8109.-01	2.0588.-01
49	2.1453.-01	6.8528.-01	9.7484.-01	2.0322.-01
50	2.1339.-01	6.8037.-01	9.6874.-01	2.0062.-01
51	2.1227.-01	6.7559.-01	9.6279.-01	1.9809.-01
52	2.1119.-01	6.7093.-01	9.5698.-01	1.9562.-01
53	2.1014.-01	6.6638.-01	9.5131.-01	1.9320.-01
54	2.0911.-01	6.6194.-01	9.4577.-01	1.9084.-01
55	2.0811.-01	6.5760.-01	9.4035.-01	1.8854.-01
56	2.0714.-01	6.5337.-01	9.3504.-01	1.8628.-01
57	2.0618.-01	6.4923.-01	9.2988.-01	1.8408.-01
58	2.0525.-01	6.4518.-01	9.2481.-01	1.8193.-01
59	2.0435.-01	6.4122.-01	9.1985.-01	1.7982.-01
60	2.0346.-01	6.3735.-01	9.1499.-01	1.7776.-01

FIGURA IV.2 bis.

ciones de factores $(K_i(1), l_i(1))$ que están realmente siendo empleadas en los distintos sectores $i = 1, N$. Puesto que en el instante inicial existe un cierto nivel de desempleo $DSMP(1) = \ell(1) - \sum_{i=1}^N l_i(1)$, es lógico tomar como límites mínimos, asimilables a cero, de las productividades marginales del trabajo, los correspondientes a dichas combinaciones.

De lo contrario, el nivel inicial de desempleo podría reducirse mediante la incorporación, todavía eficaz, de más fuerza de trabajo a algunos de los sectores.

Llamando PMM_i a las productividades marginales del factor trabajo correspondientes a la combinación inicial de factores, las restricciones $l_i(t) \leq f_i(K_i(t))$ adoptan la forma analítica :

$$\frac{\theta_i}{\ell_i} \eta_i (1 + \nu_i)^{t-1} \left[\beta_i [K_i(t)]^{-\ell_i} + (1 - \beta_i) [l_i(t)]^{-\ell_i} \right]^{-\frac{\theta_i}{\ell_i} - 1} \cdot \ell_i (1 - \beta_i) [l_i(t)]^{-\ell_i - 1} \geq PMM_i$$

Es decir, dada la dotación de capital de un sector i , ($i = 1, N$) en cualquier instante $t = 2, T$, es posible llevar los niveles de empleo solamente hasta aquel valor $l_i(t)$ para el que la productividad marginal del trabajo alcance el valor mínimo PMM_i .

Mientras subsista una situación global de desempleo, la restricción será activa en todos los sectores. Pero es preciso plantearla en la forma desigualdad porque es posible que se generen trayectorias correspondientes a estrategias de desarrollo que, absorbiendo el desempleo total, hagan que la restricción deje de ser activa en algunos sectores por lo menos.

En tal caso, el propio sistema de optimización deberá distribuir la cantidad total de trabajo entre los distintos sectores, sin que se alcancen necesariamente en todos ellos los valores PMM_i .

Insistiendo en esta idea, el valor $f_i(k_i(t))$ no representa el nivel de empleo del factor trabajo asociado a unas determinadas dotaciones sectoriales de capital, sino la máxima cantidad de trabajo que es físicamente posible emplear con tales dotaciones.

Dentro de esta cota máxima, el valor óptimo de las combinaciones de factores capital-trabajo realmente empleadas en cada instante depende de las funciones de producción y del objetivo adoptado, y será determinada en el proceso de optimización.

Los niveles de producción determinados por las funciones expuestas, se distribuyen de acuerdo con las:

5. ECUACIONES DE DISTRIBUCION

Son del tipo input-output como las utilizadas en los modelos lineales. Al representar de esta forma las exigencias en productos intermedios necesarios para llevar a cabo el proceso productivo descrito no consideramos substituciones entre los mismos. Tal hipótesis puede mantenerse con los niveles de agregación de un modelo de este tipo.

Matricialmente las expresamos como:

$$q(t) + Dq(t) + m(t) = Aq(t) + BI(t) + e(t) + c(t)$$

o bien :

$$c(t) = Pq(t) - BI(t) + m(t) - e(t)$$

donde:

- $q(t)$: vector de los niveles de producción en el periodo t .
- D : matriz de propensión marginal a importar para atender las necesidades en bienes intermedios del proceso de producción.
- $m(t)$: cantidades de bienes importados no relacionados con el proceso productivo, periodo t . Su toma en consideración explícita esta relacionada con las restricciones sobre la trayectoria que se exponen posteriormente.
- A : matriz input-output.
- B : matriz de coeficientes de capital.
(b_{ij} = cantidad del bien i necesaria para realizar una unidad de inversión a efectos de aumentar la dotación de capital del sector j)
- $I(t)$: vector de niveles de las actividades de inversión en los distintos sectores, periodo t .
- $e(t)$: vector de exportación periodo t .
- $C(t)$: cantidades destinadas al consumo, periodo t .

Observese que la división en sectores del sistema económico inducirá estructuras particulares en las matrices A y B . En particular, si la producción del sector i no contribuye a la formación de capital de ningún sector, la fila

Si una fila de la matriz A es nula, el sector correspondiente no produce bienes intermedios. Si una misma fila i es idénticamente nula en ambas matrices, el sector i solamente produce bienes de consumo final.

Las ecuaciones de distribución así especificadas, son complementadas con algunas de las siguientes

6) RESTRICCIONES SOBRE LA TRAYECTORIA

$$C_i(t) \geq C_i^*(t)$$

$$\forall t \in [1, T]$$

Estas restricciones representan la preocupación de no sacrificar, por debajo de un cierto nivel, los niveles de consumo al proceso de capitalización. En otras palabras, nos parece importante incluir entre los objetivos, la consecución de un nivel mínimo de consumo.

$$m_i(t) \geq 0$$

$$I_i(t) \geq 0$$

$$\forall t \in [1, T]$$

Implicando que no es posible descapitalizar ni efectuar importaciones con signo negativo (las exportaciones han sido tomadas en cuenta independientemente). Es de señalar que analíticamente hablando, no derivan automáticamente de la estructura del problema, debiendo ser tomadas en cuenta explícitamente

$$\sum_{i=1}^N l_i(t) \leq l(t)$$

con

$$l(t) = l(1) (1 + dP)^{t-1}$$

Restricción con la que se explicita el abandono de la hipótesis de pleno empleo. La evolución de la fuerza de trabajo se supone dada exógenamente a partir del estado inicial, con un ritmo de crecimiento dP .

Tal restricción instantánea la convertimos, a efectos de resolución del modelo, en la ecuación que define el nivel de desempleo.

$$DSMP(t) = P(t) - \sum_{i=1}^N P_i(t)$$

con $DSMP(t) \geq 0$

$$\gamma(t) \leq \gamma^*(t)$$

$$\forall t \in [1, T]$$

donde $\gamma(t)$ representa el nivel de deuda exterior en el periodo t .

Esta restricción instantánea substituye a la condición terminal dada por Kendrick $\gamma(T) = \gamma^*$. Con este planteamiento, la deuda exterior podría alcanzar cualquier nivel, por elevado que fuese, durante el periodo considerado, con tal de que tomara un valor límite γ^* al final del mismo.

Puesto que existen factores institucionales que limitan fuertemente los niveles de endeudamiento exterior de un país en desarrollo, extendemos la restricción que representa tal circunstancia a todos los periodos considerados, incluyendo al último de ellos y no solamente a este.

Una vez mas, observese la relación entre el realismo en la formulación del modelo y las posibilidades de resolución del mismo. Como es reconocido explícitamente, la adopción de tan irreales hipótesis acerca del comportamiento de los

niveles de empleo y de deuda exterior, se debe a la necesidad de evitar formulaciones que incluyan restricciones desigualdad y sus problemas computacionales asociados. La posibilidad de analizar problemas de control con tal tipo de restricciones es quien permite formular hipótesis mas realistas.

La evolución de $\delta(t)$ está ligada a la evolución del sistema mediante la ecuación de estado representada por la :

7) ECUACION DE EVOLUCION DE LA DEUDA EXTERIOR.

$$\delta(t) = (1 + IDEX) \delta(t-1) + \sum_{i=1}^N (d_{ii} q_i(t) - e_i(t) + m_i(t)) + \Gamma I(t)$$

donde:

$\delta(t)$: nivel de deuda exterior, periodo t.

IDEX : tipo de interés sobre la deuda exterior

d_{ii} : elementos diagonales de la matriz D.

Γ : vector fila de las propensiones marginales a importar ligadas a los niveles de las actividades de inversion

Esta ecuación encierra las siguientes hipótesis :

- las importaciones se dividen en tres clases : de bienes intermedios, de formación de capital y dedicados al consumo, representadas por los términos $\sum d_{ii} q_i$, $\sum m_i$, ΓI , $(i=1, N)$. La necesidad de satisfacer las

restricciones instantaneas $C_i(t) \neq C_i^*(t)$

induce a considerar las importaciones representadas por $m_i(t)$.

- los intereses sobre la deuda exterior generan nueva deuda exterior al igual que por otro concepto.
- la única forma de reducir deuda exterior es mediante el desarrollo de las actividades de exportación. Ello sería utópico y no tendría en cuenta la ayuda internacional ni otras formas de prestamo si no fuera porque no se exige en el modelo el equilibrio de la balanza exterior. En su lugar se introducen las restricciones ya enunciadas $f(t) \leq f^*(t)$.

8) ECUACIONES DE FORMACION DE CAPITAL.

Si admitiéramos la ausencia de rigideces estructurales y de rendimientos decrecientes en el proceso de formación de capital, tales ecuaciones las escribiríamos en forma vectorial:

$$K(t+1) = K(t) + I(t) - DPC \cdot K(t)$$

donde el único elemento de nueva definición es el vector linea DPC formado por los ritmos de depreciación del capital en los distintos sectores.

Ello implica suponer que:

- todos los recursos dedicados a las actividades de inversión se traducen integramente en un aumento por la misma cuantía de la dotación de capital del sector correspondiente.
- tal eficacia en la inversión se mantiene con ausencia de cualquier cota superior al nivel de I .

Ambas hipótesis son difíciles de mantener acerca de los procesos de formación de capital de una economía en vías

de desarrollo.

Numerosas rigideces institucionales, falta de preparación técnica y plazos más largos de gestación que los considerados en nuestro modelo (que no incluye retardos en el tiempo) obligan a suponer que la asimilación de la inversión es un proceso en el que :

- no es posible aumentar, en un solo periodo, la dotación de capital en cada sector por encima de una fracción μ_i del capital ya disponible :

$$\Delta k_i(t) = k_i(t+1) - k_i(t) \leq \mu_i k_i(t)$$

- el proceso de inversión hasta alcanzar tal límite superior se realiza con rendimientos decrecientes :

$$\frac{\partial \Delta k_i(t)}{\partial I_i(t)} \geq 0 \quad \frac{\partial^2 \Delta k_i(t)}{\partial^2 I_i(t)} \leq 0$$

Para representar tal fenómeno, utilizamos las funciones no lineales que representan restricciones en la asimilación de las nuevas creaciones de capital productivo propuestas por Dorfman y utilizado por Kendrick (K6).

$$g_i(k_i(t), I_i(t)) = \mu_i k_i(t) \left[1 - \left[1 + \frac{\varepsilon_i I_i(t)}{\mu_i k_i(t)} \right]^{-\frac{1}{\varepsilon_i}} \right]$$

$$\varepsilon_i \geq -1, \quad \mu_i \geq 0$$

La discusión de este tipo de función puede encontrarse en la referencia citada. Basta aquí señalar que el parámetro ε_i es una medida de la rapidez con la que decrece la eficacia de la inversión a medida que $I_i(t)$ se acerca a $\mu_i k_i(t)$, y que cuando $\varepsilon_i = -1$ la eficacia es absoluta y la relación entre inversión y

La ecuación de acumulación queda pues de la forma :

$$K(t+1) = K(t) + g(K(t), I(t)) - DPC \cdot K(t)$$

Puesto que la inclusión, en un problema de control óptimo, de variables de control bajo forma lineal en las ecuaciones de estado puede dar lugar a problemas singulares, las no-linealidades de las funciones $g(K(t), I(t))$ no presentan, antes el contrario, dificultades especiales de cálculo.

El modelo quedará completamente definido en cuanto se especifique el comportamiento o el carácter atribuido al vector $e(t)$ (exportaciones de los distintos sectores). Tales variables pueden ser considerados de distintas formas:

- a) Como elementos exogenamente determinados. Ello implica aislar de cierta forma el sistema económico del contexto internacional. El análisis de las posibles estrategias de desarrollo basadas en una actividad exportadora se limita fuertemente, a cambio de lo cual obtenemos un modelo con un menor número de variable de control.
- b) Como nuevas variables de control. En este caso se hace preciso disponer de información, o elaborar un conjunto de hipótesis, acerca del comportamiento de la demanda exterior frente a los diferentes niveles escogidos para los controles $e(t)$ dentro del conjunto factible de los mismos. En determinados casos es posible asociar los controles $e(t)$ y $m(t)$ en una única variable de decisión. Por ejemplo, si se adopta la hipótesis (evidé -
dentemente simple pero mejor que la exogeneidad de

$e(t)$, de la existencia de un límite superior no constante $e^*(t)$ en la capacidad de absorción por el mercado internacional de los bienes producidos por los distintos sectores de la economía analizada, y se admite una débil elasticidad de los precios internacionales para

valores de $e(t) \in [0, e^*(t)]$, se puede substituir $m(t) - e(t)$

por $BLC(t)$ con la condición:

$$BLC \geq -e^*(t)$$

Desde el punto de vista del método de cálculo propuesto, nada se opone a utilizar, si de ellas se dispusiera, las funciones de demanda internacional para los bienes de distintos sectores. La información por ellas facilitada determinaría la eficacia de los distintos niveles de $e(t)$ en la obtención de medios de cambio exterior y condicionaría el cálculo de sus valores óptimos. En tal caso se considerarían como variables de control independientes con:

$$e(t) \geq 0$$

Para formular el problema así expuesto como un problema de control, consideraremos, en el apartado siguiente, $e(t)$ como variables de control.

En la formulación que adoptamos, las expresiones $m_i(t) - e_i(t)$ ($i=1, N$), serán substituidas por la nueva variable de control $BLC_i(t)$, ($i=1, N$) y las restricciones $m_i(t) \geq 0$, $e_i(t) \leq e_i^*(t)$ por las restricciones $BLC_i(t) \geq -e_i^*(t)$ ($i=1, N$)

El conjunto de elementos así expuestos constituye un problema discreto de control óptimo cuyas variables de estado son $K_i(t)$, $(i=1, N)$ y $\delta(t)$, las variables de control son $l_i(t)$, $m_i(t)$, $I_i(t)$, $(e_i(t))$, $(i=1, N)$, y con las variables exógenas $l(t)$, $(e_i(t))$, $(i=1, N)$. (Según el tipo de formulación adoptada y los datos disponibles, $e_i(t)$ puede ser considerada alternativamente como variable exógena o de control). Consideremosla aquí como una variable de control.

Además de estas variables, y a fin de explicitar la estructura económica del modelo, la exposición anterior ha introducido las variables intermedias $q_i(t)$, $q(t)$, $(i=1, N)$. Sin embargo, efectuando las substituciones pertinentes, estas variables intermedias pueden ser eliminadas y todas las ecuaciones del modelo expresadas en función de las únicas variables de estado, control y exógenas arriba mencionadas.

Para ello, $q_i(t)$, $(i=1, N)$, será substituido en función de las variables de estado $K_i(t)$, y control $l_i(t)$, mediante las funciones de producción:

$$q_i(t) = \eta_i (1 + \nu_i)^t \left[\beta_i [K_i(t)]^{-\rho_i} + (1 - \beta_i) [l_i(t)]^{-\rho_i} \right]^{-\frac{\rho_i}{\rho_i}} \quad (i=1, N)$$

tanto en las ecuaciones de distribución:

$$C(t) = P q(t) - B I(t) + m(t) - e(t)$$

como en la de acumulación de deuda exterior:

$$\delta(t+1) = (1 + IDEX) \delta(t) + \sum_{i=1}^N (d_{ii} q_i(t) - e_i(t) + m_i(t)) + \Gamma I(t)$$

Asimismo, en la función objetivo se substituirán las variables intermedias $C_i(t)$, ($i=1, N$) por la expresión de las variables de estado y control resultante de las anteriores substituciones.

Con ello, el problema de control se plantea en los siguientes términos:

Dado el ESTADO INICIAL en $t=1$:

$$K_i(1) = K_i^0, \quad \delta(1) = \delta^0(1)$$

$$l_i(1) = l_i^0 \quad (i=1, N)$$

determinar las VARIABLES DE CONTROL

$$l_i(t) \quad (i=1, N), (t=2, T-1)$$

$$BLC_i(t) = m_i(t) - e_i(t) \quad (i=1, N), (t=1, T-1)$$

$$I_i(t) \quad (i=1, N), (t=1, T-1)$$

de forma que la trayectoria resultante de su aplicación a través de las ECUACIONES DE ESTADO :

$$\begin{aligned} \delta(t+1) = & (1 + IDEX) \delta(t) + \\ & + \sum_{i=1}^N \left[d_{ii} \eta_i (1 + \gamma_i)^t \left[\beta_i [K_i(t)]^{-\frac{1}{\rho_i}} + (1 - \beta_i) [l_i(t)]^{-\frac{1}{\rho_i}} + \right. \right. \\ & \left. \left. + BLC_i(t) \right] + \Gamma I(t) \right] \end{aligned}$$

$$K_i(t+1) = (1 - DPC_i) K_i(t) + \mu_i K_i(t) \left[1 - \left[1 + \frac{\epsilon_i I_i(t)}{\mu_i K_i(t)} \right]^{-\frac{1}{\epsilon_i}} \right]$$

$$(i=1, N)$$

$$(t=1, T-1)$$

y de evolución :

$$l(t+1) = l(1) (1 + \alpha P)^t \quad (t=1, T-1)$$

satisfagan las RESTRICCIONES INSTANTANEAS :

$$C_i(t) \geq c_i^*(t) \Rightarrow P_q(t) - BI(t) + BLC(t) \geq C^*(t) \quad (t=1, T-1)$$

$$0 \leq l_i(t) \leq f_i(k_i(t)) \quad (t=2, T-1), (i=1, N)$$

$$k_i(t) \geq 0 \quad (t=2, T), (i=1, N)$$

$$BLC_i(t) \geq -e_i^*(t) \quad (t=1, T-1), (i=1, N)$$

$$\delta(t) \leq \delta^*(t) \quad (t=2, T)$$

$$I_i(t) \geq 0 \quad (t=1, T-1), (i=1, N)$$

$$DSMP(t) = l(t) - \sum_{i=1}^N l_i(t) \geq 0 \quad (t=2, T-1), (i=1, N)$$

genere un ESTADO FINAL perteneciente al blanco definido por las restricciones :

$$k_i(T) \geq k_i^*$$

y de el máximo valor a la función objetivo global :

$$J = \sum_{t=1}^{T-1} \left[(1+r)^{-t} \sum_{i=1}^N a_i \left[\sum_{j=1}^N P_{ij} q_j(t) - \sum_{j=1}^N b_{ij} I_j(t) + BLC_i(t) \right]^{b_i} \right] - \sum_{t=2}^{T-1} \alpha DSMP(t) + \sum_{i=1}^N VTK_i [k_i(T) - k_i^*]$$

En todas las anteriores ecuaciones, el valor de $q_i(t)$ viene dado a través de las funciones de producción.

Particularicemos para el problema así planteado las consideraciones del capítulo III, para analizar las ventajas e inconvenientes relativos de los distintos procedimientos con los que puede abordarse su resolución.

La existencia de numerosas restricciones desigualdad del tipo $a \leq x \leq b$, sugiere el empleo de un método GRG como el expuesto en III. 3. Este método permite tratar tales restricciones de una forma eficaz, significativa (recuérdese el concepto de variables básicas y no básicas) y sin precisar la introducción de variables de holgura.

Sin embargo, el método (como todo método primal), requiere el conocimiento de una solución factible inicial, cuya búsqueda puede constituir un difícil problema (Vease en la referencia Abadie (A1), los procedimientos de búsqueda de tales soluciones).

La resolución por el método dual de los multiplicadores expuesto en III. 4. 2 presenta un cuadro de ventajas e inconvenientes simétrico del anterior. Es decir, no es preciso conocer una solución factible de partida pero si es posible convertir en igualdades todas las restricciones del problema, con el consiguiente número de variables de holgura suplementarias así generadas.

Sin embargo, dada la forma en que se presenta el problema estudiado, la introducción de variables de holgura positivas puede ir acompañada de un proceso de substitución y eliminación de variables que permita no aumentar excesivamente el número de las mismas.

Por otra parte, el método dual de los multiplicadores

(como todo método dual con respecto a un método primal), tiene a su favor la mayor simplicidad numérica de su proceso de cálculo. En particular, la solución del problema de control aquí planteado no requiere mas que una serie de multiplicaciones vectoriales y un proceso de búsqueda unidimensional. En cambio, el método GRG propuesto, a pesar de su estructura escalonada en el cálculo del gradiente reducido, requiere la solución de sistemas de ecuaciones no lineales e inversiones de matrices.

El anterior planteamiento del problema precisa de muy pocas modificaciones para su resolución mediante el método GRG. Concretamente basta con introducir variables de holgura positivas para convertir en igualdades las restricciones desigualdad donde intervengan más de una variable. De acuerdo con ello, las restricciones:

$$C_i(t) \geq C_i^*(t) \quad (i=1, N) \quad (t=1, T-1)$$

$$l_i(t) \leq f_i(k_i(t)) \quad (i=1, N) \quad (t=2, T-2)$$

$$l(t) - \sum_{i=1}^N l_i(t) \geq 0 \quad t=2, T-1$$

se convierten en :

$$C_i(t) - H C_i(t) = C_i^*(t)$$

$$l_i(t) + H l_i(t) = f_i(k_i(t))$$

$$l(t) - \sum_{i=1}^N l_i(t) - DSMP(t) = 0$$

con las restricciones adicionales del tipo $a \leq x \leq b$:

$$H C_i(t) \geq 0 \quad (i=1, N) , (t=1, T-1)$$

$$H l_i(t) \geq 0 \quad (i=1, N) , (t=2, T-1)$$

$$DSMP(t) \geq 0 \quad (t=2, T-2)$$

Las relaciones $l_i(t) + H l_i(t) = f_i(k_i(t))$ permiten la eliminación inmediata de las variables $l_i(t)$ que serán substituidas en todas las expresiones donde aparezcan por $f_i(k_i(t)) - H l_i(t)$ y su valor calculado a posteriori a través de la anterior igualdad. Evidentemente, las demás igualdades podrían también utilizarse para eliminar otras variables, pero el proceso no siendo tan inmediato, pasemos a considerar el número total de variables y restricciones, otras que las del tipo $a \leq x \leq b$, que aparecerían en la solución por el método GRG:

NUMERO DE VARIABLES:

$K_i(t)$	$\begin{cases} i=1, N \\ t=2, T \end{cases}$
$\delta(t)$	$t=2, T$
$I_i(t)$	$\begin{cases} i=1, N \\ t=1, T-1 \end{cases}$
$BLC_i(t)$	$\begin{cases} i=1, N \\ t=1, T-1 \end{cases}$
$H l_i(t)$	$\begin{cases} i=1, N \\ t=2, T-1 \end{cases}$
$DSMP(t)$	$t=2, T-1$
$HC_i(t)$	$\begin{cases} i=1, N \\ t=1, T-1 \end{cases}$

Parametrizado en función del número de sectores y del número de periodos de tiempo, ello equivale a un total de $5NT - 6N + 2T - 3$ variables.

NUMERO DE RESTRICCIONES OTRAS QUE $a \leq x \leq b$

, Ecuaciones de estado : $(N+1)(T-1)$
 Restricciones de consumo mínimo : $N(T-1)$
 Restricciones de utilización del factor trabajo : $(T-2)$

Lo que equivale a un total de $2NT - 2N + 2T - 3$

Para fijar ideas, un plan extendido a un horizonte de 20 periodos de tiempo y un nivel de desagregación de cuatro sectores implicaría la resolución de un modelo con 413 variables y 189 restricciones otras que las del tipo $a \leq x \leq b$. Extenderlo a 30 periodos implicaría la consideración de 633 variables y 289 restricciones.

Por el contrario, un intento de solución por el método dual de los multiplicadores exige una reformulación y adaptación previa más laboriosa: Este es el objeto del siguiente apartado.

IV. 3. 3 RESOLUCION POR EL METODO DUAL DE LOS MULTIPLICADORES.

Como en el anterior planteamiento, el conocimiento del estado del sistema permite dejar de considerar como tales a las siguientes variables:

$l_i(1)$

$k_i(1) \quad (i=1, M)$

$\delta(1)$

$DSMP(1)$

que quedan fijadas en los valores tomadas como datos.

Para evitar una restricción explícita $(x \geq 0)$ de positividad sobre las variables de holgura que se introduzcan,

estas se expresaran ahora como cuadrados de variables.

Sin embargo, a fin de limitar el número de variables de holgura a introducir, algunas de las restricciones del tipo $a \leq x \leq b$ pueden ser tomadas en cuenta implícitamente. En efecto, así ocurre con las restricciones de positividad en los factores de producción:

$$\begin{array}{ll} l_i(t) \geq 0 & k_i(t) \geq 0 \\ (t=2, T-1) & (t=2, T) \end{array}$$

que pueden ser consideradas mediante la definición de las funciones de producción de forma tal que generen niveles de producción nulos cuando uno, u ambos, de los factores, tome valores negativos.

Por su parte, las restricciones :

$$\left. \begin{array}{l} C_i(t) = \sum_{j=1}^N P_{ij} q_j(t) - \sum_{j=1}^N b_{ij} I_j(t) + BLC_i(t) \geq C_i^*(t) \\ I_i(t) \geq 0 \\ \left. \begin{array}{l} l_i(t) \leq f_i(k_i(t)) \\ l(t) - \sum_{i=1}^N l_i(t) \geq 0 \end{array} \right\} t=2, T-1 \\ BLC_i(t) \geq -e^*(t) \quad t=1, T-1 \\ \delta(t) \leq \delta^*(t) \quad t=2, T \\ K_i(T) \geq K_i^* \end{array} \right\} i=1, N$$

se convertirán en igualdades mediante variables de holgura positivas que permitan, a su vez, la eliminación de parte de ellas. Una vez introducidas dichas variables, las ecuacio-

nes resultantes son :

$$1 \quad \sum_{j=1}^N P_{Lj} q_j(t) - \sum_{j=1}^N b_{ij} I_j(t) + BLC_i(t) - HC_i^2(t) = C_i^*(t)$$

$$2 \quad I_i(t) - HI_i^2(t) = 0$$

$$3 \quad l_i(t) + Hl_i^2(t) = f_i(k_i(t))$$

$$4 \quad l(t) - \sum_{i=1}^N l_i(t) - HDSMP^2(t) = 0$$

$$5 \quad BLC_i(t) - HBLC_i^2(t) = -e_i^*(t)$$

$$6 \quad \delta(t) + H\delta^2(t) = \delta^*(t)$$

$$7 \quad K_i(T) - HK_T i^2 = K_i^*$$

Las ecuaciones 2,3,5,6 y 7 permiten la eliminación inmediata de las variables $I_i(t)$, $l_i(t)$, $BLC_i(t)$ y $\delta(t)$ respectivamente. En aquellas ecuaciones del modelo donde aparezcan, estas variables serán substituidas por las expresiones derivadas de las anteriores ecuaciones :

$$I_i(t) = HI_i^2(t)$$

$$l_i(t) = f_i(k_i(t)) - Hl_i^2(t)$$

$$BLC_i(t) = HBLC_i^2(t) - e_i^*(t)$$

$$\delta(t) = \delta^*(t) - H\delta^2(t)$$

$$K_i(T) = K_i^* + HK_T i^2$$

y su valor calculado "a posteriori" mediante las mismas e-

cuaciones.

En particular, las ecuaciones 1 y 4 quedan, después de efectuar dichas substituciones, de la forma:

$$1) \sum_{j=1}^N P_{ij} q_j(t) - \sum_{j=1}^N b_{ij} H I_j^2(t) + H B L C_i^2(t) - e_i^*(t) - C_i^*(t) = H C_i^2(t)$$

$$4) \ell(t) - \sum_{i=1}^N (f_i(K_i(t)) - H \ell_i^2(t)) = H D S M P^2(t)$$

Sin efectuar las substituciones y eliminaciones que no sean tan inmediatas como las citadas, el problema queda definido en función de las únicas variables:

$$\left. \begin{array}{l} H C_i(t) \\ H I_i(t) \\ H B L C_i(t) \end{array} \right\} \begin{array}{l} t = 1, T-1 \\ i = 1, N \end{array}$$

$$\left. \begin{array}{l} H \ell_i(t) \\ K_i(t) \end{array} \right\} \begin{array}{l} t = 2, T-1 \\ i = 1, N \end{array}$$

$$H D S M P(t) \quad t = 2, T-1$$

$$H \ell(t) \quad t = 2, T$$

$$H K T_i \quad i = 1, N$$

como el de :

$$\text{Max } J =$$

$$= \sum_{t=1}^{T-1} \left[(1+\rho)^{-t} \sum_{i=1}^N a_i \left[C_i^*(t) + H C_i^2(t) \right] b_i \right] - \sum_{t=2}^{T-1} \alpha H D S M P^2(t) + \sum_{i=1}^N V T K_i \cdot H K T_i^2$$

bajo las restricciones :

$$1) C_i^*(t) + H C_i^2(t) = \sum_{j=1}^N p_{ij} q_j(t) - \sum_{j=1}^N b_{ij} H I_j^2(t) + H B L C_i^2(t) - e_i^*(t)$$

$$(i = 1, N)$$

$$(t = 1, T-1)$$

en la que :

$$q_j(t) = \eta_j (1 + \nu_j)^t \left[\beta_j [K_j(t)]^{-\frac{p_j}{\beta_j}} + (1 - \beta_j) [f_j(K_j(t)) - H \ell_j^2(t)]^{-\frac{p_j}{\beta_j}} \right]^{-\frac{\beta_j}{p_j}}$$

$$\text{si } K_j(t) \geq 0 \quad \gamma \quad f_j(K_j(t)) - H \ell_j^2(t) \geq 0$$

$$q_j(t) = 0$$

$$\text{si } K_j(t) < 0 \quad \delta \quad f_j(K_j(t)) - H \ell_j^2(t) < 0$$

$$2) K_i(t+1) = (1 - D P C_i) K_i(t) + \mu_i K_i(t) \left[1 - \left[1 + \frac{\varepsilon_i H I_i^2(t)}{\mu_i K_i(t)} \right]^{-\frac{1}{\varepsilon_i}} \right]$$

$$(i = 1, N)$$

$$(t = 1, T-2)$$

$$H K T_i^2 = (1 - D P C_i) K_i(T-1) + \mu_i K_i(T-1) \left[1 - \left[1 + \frac{\varepsilon_i H I_i^2(T-1)}{\mu_i K_i(T-1)} \right]^{-\frac{1}{\varepsilon_i}} \right] - K_i^*$$

$$(i = 1, N)$$

$$3) \delta^*(t+1) - H \delta^2(t+1) = (1 + |D E X|) (\delta^*(t) - H \delta^2(t)) + \sum_{i=1}^N [d_{ii} q_i(t) + H B L C_i^2(t) - e_i^*(t)] + \Gamma H I^2(t)$$

$$(t = 1, T-1)$$

$$4) H D S M P^2(t) = \ell(t) - \sum_{i=1}^N [f_i(K_i(t)) - H \ell_i^2(t)]$$

$$(t = 2, T-1)$$

El número total de variables y de restricciones (todas las existentes son ahora del tipo igualdad) que contiene esta formulación es:

NUMERO DE VARIABLES: $5TN - 6N + 2T - 3$

NUMERO DE RESTRICCIONES: $2NT - 2N + 2T - 3$

Tales resultados muestran que, mediante los procedimientos de substitución y eliminación empleados, el número de variables y de restricciones a considerar son los mismos tanto si se utiliza un procedimiento GRG como el método dual de los multiplicadores. En el caso de GRG habrían restricciones adicionales de la forma $a \leq x \leq b$ que son tomados directamente en consideración por el proceso de cálculo, mientras que, de emplearse el método dual de los multiplicadores, es preciso efectuar una etapa final de cálculo para deducir los valores de las verdaderas variables del problema a través de las ficticias empleadas en la resolución.

Se habrá observado que, ni un método GRG ni el método dual de los multiplicadores, producen ningún tipo de descomposición estructural del problema.

Ciertamente, el método GRG aprovecha la estructura del Jacobiano de las restricciones, para calcular secuencialmente las componentes del gradiente reducido. Pero la solución se obtiene de forma global al calcularse conjuntamente los componentes del vector x .

Por su parte, el método dual de los multiplicadores, aunque radicalmente distinto en su concepción y proceso de cálculo, no llega a producir una descomposición de los problemas del primer nivel en subproblemas locales. El proceso de modificación del vector primal se realiza conjuntamente en todas sus componentes y la solución sigue presentando un carácter global, a pesar de la inspiración dual del algoritmo.

Hemos visto el elevado número de variables y restricciones que nos vemos obligados a considerar simultáneamente en el empleo de tales métodos. Cuando Kendrick y Taylor (K6) resuelven (en 1970) un modelo sin restricciones con cuatro sectores y 30 periodos de tiempo, apuntaban la posibilidad de resolver mediante tales procedimientos problemas de un nivel de desagregación de hasta 10 sectores. Abadie y Robert (A1), al resolver en 1973 un problema con el mismo número de sectores y periodos, recogen tal posibilidad aunque la complementan con un cauteloso comentario: "...la considération de problèmes ayant, disons, 10 secteurs, est parfaite-

ment envisageable,... Il ne faut pas se dissimuler cependant qu'il s'agit d'un problème difficile. Avons nous même atteint la solution optimale du problème à 4 secteurs ? ".

El método dual de los multiplicadores, al no efectuar inversiones ni soluciones de sistemas de ecuaciones, puede ser, quizá, menos sensible al aumento del número de variables. Pero la consideración de niveles de desagregación que no sean simplemente académicos, producen números de variables y restricciones demasiado grandes para que un método global de resolución pueda ser indefinidamente aplicable.

Estas consideraciones, que son de naturaleza exclusivamente técnica o numérica, demuestran la necesidad algorítmica de un proceso de descomposición que permita substituir el problema global por una serie de problemas de menores dimensiones.

Si del proceso de resolución, pasamos a considerar la estructura económica del modelo, la solución global del mismo aparece como igualmente poco deseable. En efecto, todo intento de solución de este tipo representa hipótesis irreales, y es, conceptualmente poco ilustrativo.

Desde un punto de vista estructural, la autoridad decisoria no puede suponerse concentrada en las manos de un solo centro. La funcional objetivo, caso de considerarse como única, no es sino la agregación de un serie de objetivos particulares de los distintos subsistemas que sea posible definir en la estructura del modelo. Tales subsistemas estando evidentemente relacionados entre si por un conjunto de restricciones comunes.

. Ello indica la necesidad conceptual de descentralizar el proceso de optimización. Desde esta óptica, el problema que se plantea es el de cómo resolver (si ello es posible) el problema global, dejando que los subsistemas (centros locales de decisión) resuelvan sus problemas locales mediante la adopción de decisiones relativas únicamente a su propio funcionamiento, de forma tal que se satisfagan las restricciones globales que los interrelacionan.

Ambas series de consideraciones, la algorítmica y la conceptual, conducen pues a las mismas conclusiones bajo los nombres alternativos de descomposición-descentralización. El primero aparece así como la traducción numérica de una realidad estructural, o el segundo como la interpretación económica de un proceso de cálculo.

En cualquier caso, lo expuesto nos conduce a considerar necesariamente varios centros de decisión, asociados a la descomposición en subproblemas que la estructura del modelo permita realizar, y al abandono de los métodos globales de resolución.

Se observará que esto es lo que pretende realizar el algoritmo dual del capítulo III. 4. 3. La diferencia con el enfoque aquí expuesto es que la única estructura que tal algoritmo explota es la propia estructura dinámica del problema de control, los subproblemas considerados siendo los ficticios asociados con cada instante discretizado del tiempo.

Al analizar problemas concretos, es posible descubrir estructuras más ricas, que permitan considerar subsistemas

económicamente más significativos y descomponer el proceso de cálculo en niveles múltiples.

Este tema es desarrollado en el siguiente capítulo, donde los métodos jerárquicos de control son expuestos y empleados en un planteamiento multinivel del problema de Kendrick y Taylor modificado. Asimismo, recogemos y desarrollamos las consideraciones del capítulo III acerca de las condiciones de aplicación de tales métodos y el valor de la información por ellos generada.

DESCENTRALIZACION ESTRUCTURAL POR DESCOMPOSICION JERARQUICA
DEL PROBLEMA DE CONTROL

"... The combination of centralized management and economic independence of the parts, constitutes the essential road towards the improvement of the national economies".

Fedorenko. (F2).

- V. 1 INTRODUCCION. LOS SISTEMAS MULTINIVELES DE CONTROL.
- V. 2 ALGORITMOS JERARQUICOS DE CONTROL OPTIMO.
- V. 3 PLANTEAMIENTO DESCENTRALIZADO DEL MODELO DINAMICO DE PLANIFICACION.
- V. 4 POSIBILIDADES, LIMITES E INTERES DE LA APLICACION DEL METODO.

De todas las hipótesis simplificadoras que encierra el planteamiento del modelo de planificación multisectorial expuesto, la más irreal e indeseable es la de la existencia de un centro de decisión único del que dependen todas las variables del sistema económico.

Dos líneas fundamentales de desarrollo pueden seguir los intentos de introducir centros de decisión múltiples en problemas dinámicos de planificación.

La primera de ellas es considerarlos como juegos diferenciales. Cada uno de los centros de decisión inmersos e interrelacionados en la estructura del sistema global posee una función objetivo propia y sus intereses están, o al menos así lo creen, en relativa contraposición con los de los demás centros. El tipo de soluciones que pueden buscarse al problema depende de la actitud supuesta de los centros de decisión y suele consistir generalmente en el cálculo de trayectorias de equilibrio de Nash o, si los centros adoptan una actitud cooperativa, en la búsqueda de trayectorias de Pareto.

La más interesante realización en esta línea es sin duda el trabajo de Pau (P3) presentado en la IFAC/IFORS International Conference on Dynamic Modelling and Control of National Economics de Julio de 1973. En este trabajo, calificado de "beautiful" por Kendrick en (A10), Pau calcula las trayectorias de Nash de un modelo tipo Kendrick de 4 sectores de la economía danesa, en el que cada uno de los sectores es considerado como un centro de decisión con una función objetivo propia. Se trata sin duda de una línea de

desarrollo de extraordinario interés, a pesar de las numerosas dificultades de cálculo asociadas. Las referencias (K12 , V2) constituyen otros ejemplos de la misma.

Podríamos decir que la óptica no cooperativa que suponen estos modelos los hace más propios de un intento de modelización de una economía capitalista. En modelos destinados a, o que parten del supuesto de, economías planificadas, no parece tener mucho sentido que la contraposición de intereses de las partes integrantes del sistema no haya sido objeto de una negociación (u imposición) política previa que haya permitido definir un objetivo común a todo el sistema a partir de los objetivos de las partes. Es decir, en una economía socialista planificada es posible suponer que las partes han constituido una "entente" y que el objetivo global se concibe como una combinación aditiva de los intereses respectivos de cada una de ellas.

En este sentido, la búsqueda de trayectorias de Nash no presenta un interés especial y la línea "juegos diferenciales" cede la importancia a los métodos de descomposición jerárquica o de control a niveles múltiples.

Tales métodos, están siendo desarrollados para la resolución práctica de problemas de control óptimo de sistemas dinámicos de elevada dimensionalidad que esten constituidos por subsistemas interrelacionados. Muchos de los sistemas dinámicos encontrados en la práctica poseen este tipo de estructura, que esquematiza la figura nº FV.1

Los subsistemas están interconectados porque los outputs de cada uno de ellos son, en parte, inputs de los demás. Cada subsistema posee sus inputs o controles propios

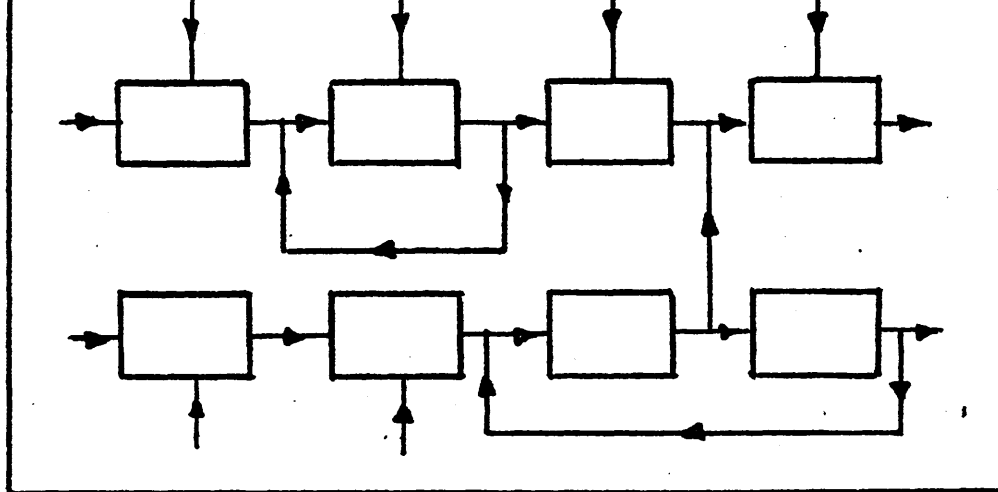


FIGURA FV.1

que le llegan independientemente de los demás subsistemas, y parte de sus outputs se dirigen al exterior del sistema global donde, presumiblemente, presentan un determinado valor. Esta situación es la representada en la figura FV.2

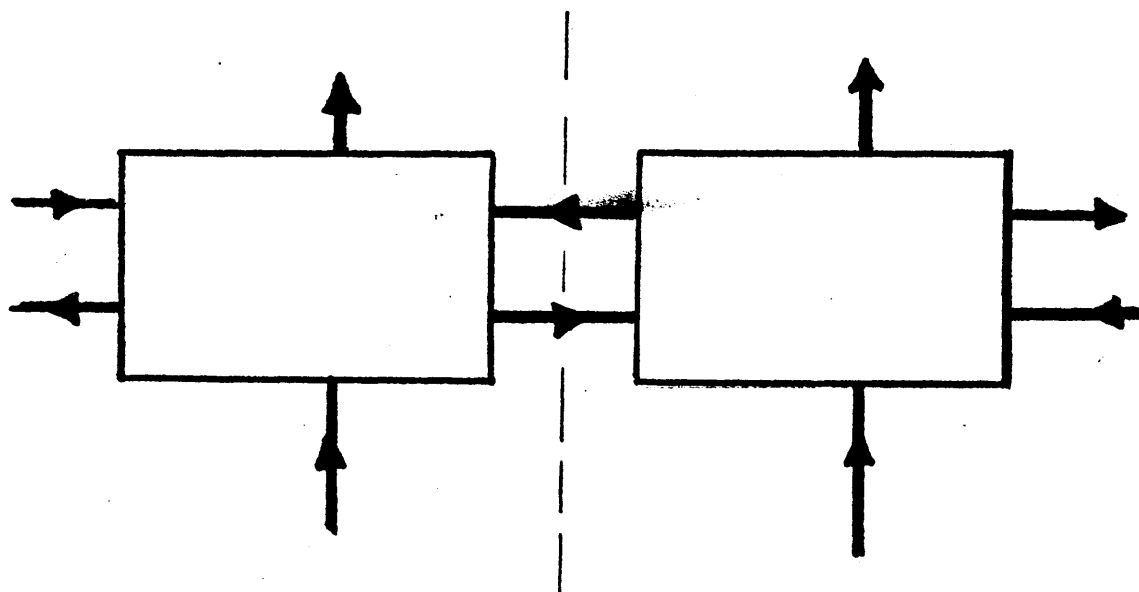


FIGURA FV.2

La descomposición en subsistemas interrelacionados induce a imaginar una descomposición del proceso de control del sistema global, a través del establecimiento de unas unidades autónomas responsables de las distintas partes, que

actuen bajo la supervisión o coordinación de un centro jerárquicamente superior que sea consciente de las interrelaciones existentes entre ellas.

Evidentemente, la descomposición del problema en niveles jerárquicos no tiene porque limitarse a dos de ellos. Un esquema idealizado con tres niveles de control de un sistema es el representado por la figura FV.3

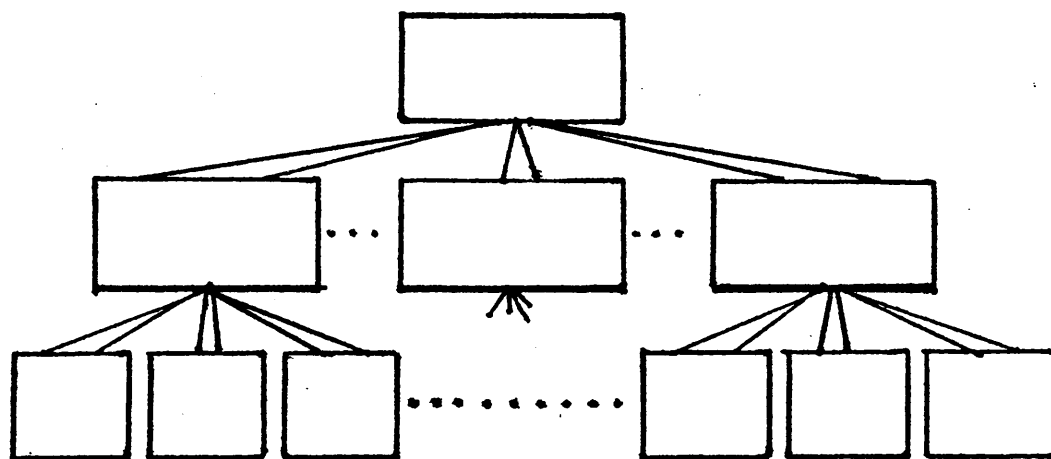


FIGURA FV.3

Para una exposición detallada de la descomposición de sistemas en estructuras jerárquicas, es preciso ver la obra básica de Mesarovic, Macko y Takahara (M7) que suponemos conocida del lector en sus líneas fundamentales.

Intuitivamente surge la idea de asimilar los subsistemas interrelacionados que componen un sistema económico nacional a los distintos sectores en que se desagrega tradicionalmente la actividad del conjunto.

Los métodos jerárquicos de control a niveles múltiples nos interesan especialmente porque pueden permitir asociar la necesidad numérica de descomposición con el proceso de descentralización correspondiente al reconocimiento de la existencia en el sistema económico de centros locales de de-

cisión dotados de amplios niveles de autonomía.

Algunos de los métodos de control óptimo por descomposición jerárquica están íntimamente relacionados con los métodos duales de descomposición en programación matemática. En el siguiente apartado de este capítulo exponemos el proceso algorítmico de uno de los más interesantes métodos de descomposición jerárquica, perteneciente al tipo de los denominados "métodos infactibles". En el apartado V. 3, lo aplicamos a un planteamiento descentralizado del modelo de planificación dinámica del capítulo IV.

Es preciso distinguir claramente la diferencia entre los métodos de descomposición jerárquica a los que nos referimos y a otros tipos de algoritmos que usan sistemas de precios en el mecanismo de descomposición-coordinación.

El método de descomposición de Dantzig y Wolf, por ejemplo, no descentraliza completamente el proceso de solución, puesto que la responsabilidad final del proceso de decisión corresponde a la autoridad central ("master program"). Este determina la solución óptima del problema conjunto mediante combinaciones convexas de las propuestas presentadas por los subprogramas. Si bien determina en cada iteración nuevos sistemas de precios que son remitidos a los subproblemas, estos solamente tienen el derecho de elaborar propuestas de solución sin que tengan que ser necesariamente incluidas en la solución global que es elaborada por el programa principal.

Los métodos jerárquicos del tipo que aquí exponemos pretenden completar la descentralización aumentando la autoridad de los centros responsables de las partes. La de-

terminación de los valores óptimos de las variables de decisión les corresponderá a ellas, a cada una dentro de su propia esfera de actividad. La autoridad central efectuará una labor coordinadora que solamente se traducirá en generación de sistemas de precios de una pequeña parte del número total de variables del sistema.

El método está dotado de una rica interpretación económica que exponemos en los dos apartados siguientes, pero dado lo ambicioso de su planteamiento, no es de extrañar que no sea aplicable a todos los casos.

Solamente lo es cuando el Lagrangiano del problema global posee un punto de silla. Sabemos que la convexidad del problema es una condición suficiente para garantizar su existencia y por lo tanto la aplicabilidad del método. En el punto V. 4 analizamos las condiciones de aplicación más allá de las condiciones suficientes representadas por la convexidad. Por otra parte, de acuerdo con los razonamientos expuestos en el capítulo III, analizamos el valor de la información generada por el empleo de tales métodos aunque de su aplicación no se obtenga la solución al problema específicamente planteado.

Una creciente literatura desarrolla algoritmos jerárquicos de control, tanto para el caso discreto como el continuo. Los métodos propuestos pueden clasificarse en tres grandes categorías; las dos primeras de las cuales son solamente aplicables a problemas con objetivo global separable :

1) Métodos "factibles" ("feasible methods")

Se caracterizan (y a ello deben su denominación), porque los sistemas de control generados en cada iteración, satisfacen todas las restricciones del problema. Pero su aplicación está limitada a problemas que poseen un número de variables de control mayor o igual que el número de interconexiones. Una completa exposición de tales métodos es la dada por Pearson (p4)

2) Métodos "infactibles" ("infeasible methods")

Con ellos, las restricciones globales solo son satisfechas por la solución óptima, por lo que no es posible utilizar los sistemas de control generados en las iteraciones previas.

3) Métodos mixtos (Titli (g4 , g5))

Son especialmente útiles en su aplicación a problemas cuyo objetivo global no es separable. Tampoco consiguen la satisfacción de todas las restricciones hasta que el óptimo es alcanzado. Pueden pues, ser considerados como una subclase de los métodos infactibles.

Limitándonos al caso de objetivos separables, los méto-

dos infactibles son especialmente interesantes por lo económicamente ilustrativo que resulta su proceso de cálculo y por el papel que en ellos juega la dualidad.

Entre ellos, los más interesantes por su aplicabilidad práctica, son los del tipo Lasdon-Tamura (13, 15). Como ya hemos indicado, el algoritmo dual del punto III. 4. 3 es un algoritmo Lasdon-Tamura que considera al problema de control dotado exclusivamente de su propia estructura dinámica.

En problemas que poseen estructuras específicas, es posible efectuar descomposiciones más ilustrativas. Consideremos de nuevo el problema discreto de control óptimo cuya formulación general es la expuesta en el Apéndice II. Supongamos que, por la forma de su objetivo y restricciones, es posible considerar divididos los vectores de estado y de control en S grupos de componentes:

$$X(t) = \begin{bmatrix} x_1(t) \\ \vdots \\ x_i(t) \\ \vdots \\ x_s(t) \end{bmatrix}, x_i \in E^{n_i} \quad \mu(t) = \begin{bmatrix} \mu_1(t) \\ \vdots \\ \mu_i(t) \\ \vdots \\ \mu_s(t) \end{bmatrix}, \mu_i \in E^{m_i}$$

tales que es posible detallar su enunciado de la siguiente forma :

$$\text{Min } J = \sum_{i=1}^S \Phi_i(x_i(T)) + \sum_{i=1}^S \sum_{t=0}^{T-1} f_{0,i}(x_i(t), \mu_i(t), t)$$

con las ecuaciones de estado:

$$\begin{aligned} i=1, S \quad x_i(t+1) &= f_i(x_i(t), \mu_i(t), t) \quad (t=0, 1, \dots, T-1) \\ x_i(0) &= x_i^0 \end{aligned} \quad (1)$$

y bajo los dos tipos de restricciones:

$$\begin{aligned} \text{I) } i=1, S \quad g_i(x_i(t), \mu_i(t), t) &\leq 0 \quad t=0, \dots, T-1 \end{aligned} \quad (2)$$

$$g_i(x_i(T)) \leq 0$$

$$\text{II)} \quad \sum_{i=1}^S h_i(x_i(t), u_i(t), t) \leq 0 \quad t=0, \dots, T-1 \quad (3)$$

Se observa inmediatamente que, de no ser por las restricciones del tipo II), el problema es en realidad la superposición de S problemas independientes de control óptimo discreto. Ni en la función objetivo de cada uno de ellos:

$$\text{Min } J_i = \Phi_i(x_i(T)) + \sum_{t=0}^{T-1} f_{0,i}(x_i(t), u_i(t), t)$$

ni en las ecuaciones (1), (2) representativas de su dinámica y restricciones específicas, intervienen las variables de estado o de control de los demás problemas.

La consideración global del problema es debida, única y exclusivamente, a las restricciones del tipo (3), que desde ahora denominamos restricciones de interconexión. Obsérvese que la estructura supuesta no solamente posee un objetivo separable, sino que también son separables aditivamente los componentes de las restricciones de interconexión.

Suponiendo que el número de estas sea RI, los S vectores $h_i, (i=1, S)$ son vectores de funciones con RI componentes en cada uno de los cuales solamente intervienen las variables específicas del grupo i.

El Lagrangiano del problema global, referido a las únicas restricciones de interconexión es :

$$\begin{aligned} \mathcal{L}(x, u, \lambda) = & \sum_{i=1}^S \left[\Phi_i(x_i(T)) + \sum_{t=0}^{T-1} f_{0,i}(x_i(t), u_i(t), t) \right] + \\ & + \sum_{t=0}^{T-1} \left[\lambda'(t) \sum_{i=1}^S h_i(x_i(t), u_i(t), t) \right] \quad (4) \end{aligned}$$

donde $\lambda(k)$ es el vector de multiplicadores de dimensión RI, asociado en cada instante k a las ecuaciones de restricción.

La minimización del Lagrangiano global con respecto a los vectores X, u , se descompone, sin más que reconocer visualmente su estructura aditiva, en la minimización de S Lagrangianos independientes, cada uno de ellos sometido a sus restricciones específicas (1), (2), que no han sido tenidas en cuenta en la formulación del Lagrangiano global y expresado en sus únicas variables:

$$\begin{aligned} X_i(t), t=1, T \\ u_i(t), t=0, T-1 \end{aligned}$$

Los S subproblemas son:

$$\text{Para } i = 1, S \\ \text{Min } \Phi_i(X_i(T)) + \sum_{t=0}^{T-1} \left[f_0(X_i(t), u_i(t), t) + \lambda'(t) h_i(X_i(t), u_i(t), t) \right]$$

bajo las restricciones:

$$X_i(t+1) = f_i(X_i(t), u_i(t), t) \quad (t=0, T-1)$$

$$X_i(0) = X_i^0$$

$$g_i(X_i(t), u_i(t), t) \leq 0 \quad (t=0, T-1)$$

$$g_i^T(X_i(T)) \leq 0$$

El cálculo de la función dual

$$\omega(\lambda) = \min_{X, u} \left\{ \mathcal{L}(X, u, \lambda) \right\} \\ \left\{ \begin{array}{l} X, u \text{ satisfaciendo} \\ (1) \quad , \quad (2) \end{array} \right\}$$

se ve extraordinariamente facilitado por la anterior descomposición de los problemas del primer nivel. La función del segundo nivel es el de facilitar una trayectoria inicial

$\lambda(t)$ de las variables duales asociadas a las restricciones de interconexión y de modificar secuencialmente dicha

trayectoria para maximizar la función dual.

El problema del segundo nivel es pues:

$$\begin{array}{c} \text{Max } \omega(\lambda) \\ \lambda \geq 0 \end{array}$$

y el esquema del algoritmo a dos niveles, representado en la figura nº FV.4 es el siguiente :

- a) Definición de una trayectoria inicial de las variables duales $\lambda(\cdot)$:

$$\lambda^0(k) , k = 0, T-1$$

- b) Resolución por los centros de decisión locales del primer nivel, de S problemas independientes de control óptimo cuya función objetivo es el Lagrangiano local correspondiente.

- c) Cálculo de la función dual por simple substitución de las soluciones de los problemas locales en la expresión (4)

- d) Maximización de la función dual por la autoridad coordinadora del segundo nivel y consiguiente modificación de la trayectoria de las variables duales.

- e) Máximo alcanzado ?

SI. FIN del proceso iterativo.

NO. Vuelta a b) utilizando la anterior solución de cada subproblema como punto inicial de las iteraciones para encontrar su nueva solución.

La aplicación práctica del método puede realizarse de

distintas formas, dependiendo de como se efectue la resolución de los problemas locales del primer nivel y el problema del segundo nivel.

Puesto que la dimensión de los problemas del primer nivel es muy reducida comparada con las dimensiones iniciales del problema, cada uno de ellos puede tratar de resolverse mediante un método primal como el GRG o utilizando el método dual de los multiplicadores expuesto. Pero es posible tratar de asociar a la descomposición estructural efectuada, una nueva descomposición temporal de los problemas del primer nivel mediante la aplicación a cada uno de ellos del Algoritmo dual Lasdon-Tamura de III. 4. 3.

El resultado sería un algoritmo a tres niveles como el representado en la figura nº FV.5 . En el, cada centro de decisión del segundo nivel resuelve su problema local bajo las condiciones impuestas por el centro de coordinación del tercer nivel a través de las variables $\lambda(k)$, $k=0, T-1$. Para así hacerlo, se constituye a su vez como centro coordinador de las subunidades temporales responsables de la optimización de sus Hamiltonianos instantáneos.

En el flujo general de información del proceso a tres niveles de la figura nº FV.5 , el lector reconocerá en cada uno de los elementos jerárquicos que tienen por vértice superior un centro de decisión del segundo nivel (como el encerrado en trazos), a los representados más detalladamente en la figura nº FIIL.8 del capítulo III. Sin repetir aquí de nuevo, con el detalle ya tratado en ese capítulo, el proceso de optimización del segundo nivel, pasemos a considerar el de maximización de la función dual. Estas consideraciones se basan fundamentalmente en las referencias (Lasdon (L3),

Dos tipos de métodos distintos son básicamente utilizables :

a) Los basados en la diferenciabilidad de la función dual.

La función dual es fácilmente calculable y sabemos que posee la propiedad (L5) de ser cóncava sobre cualquier subconjunto convexo de su dominio de definición, independientemente de la convexidad del problema. Si el gradiente de la función dual con respecto a sus variables $\lambda(k)$ pudiera ser fácilmente obtenido, estarían reunidas las condiciones óptimas para la aplicación de un método del tipo gradiente, convenientemente modificado para tener en cuenta las restricciones .

En general, la función dual no es diferenciable en todos los puntos de su dominio de definición, ni siquiera en problemas convexos. (Para un estudio detallado de las condiciones de diferenciabilidad de la función dual, vease Lasdon (L3)). Si la función dual se considera diferenciable en todos los puntos de interés, se obtiene (L3) un resultado fundamental : las derivadas parciales de la función dual con respecto a las variables $\lambda(k)$ son las que se obtendrían si se derivara el Lagrangiano sin tener en cuenta la dependencia de las variables primales $x(k)$, $u(k)$ con respecto a las variables $\lambda(k)$. Es decir, efectuando tal operación y aplicando este resultado se obtiene :

$$\nabla_{\lambda(k)} \omega(\lambda) = \sum_{i=1}^S h_i(x_i(k), u_i(k), \lambda) , (k=0, T-1)$$

En otras palabras, el gradiente de la función dual es igual al valor que tomen las restricciones de inter-

conexión como consecuencia de las decisiones elaboradas por los centros del 1^{er} nivel (2º nivel si consideráramos la descomposición en 3 niveles).

b) Los métodos de aproximación tangencial.

Tales métodos, propuestos por Geoffrion (G2) y menos desarrollados en la práctica de lo que su atractiva concepción teórica merece, no precisan la diferenciabilidad de la función dual y evitan las dificultades en la convergencia de los anteriores métodos cuando la función dual posee numerosos puntos con derivadas discontinuas.

Tales procedimientos se basan en la concavidad de la función dual y en el hecho de que, al minimizar el Lagrangiano global para un valor dado de las variables $\lambda(\cdot)$, $\lambda(\cdot) = \lambda^*(\cdot)$, se obtiene inmediatamente un hiperplano soporte de la función dual en el punto $\lambda^*(\cdot)$. Gracias a ello, la maximización de la función dual se substituye por la resolución de un programa lineal al que se le van añadiendo restricciones sucesivamente.

La exposición detallada de este tipo de métodos sería demasiado extensa y nos alejaría de nuestro interés básico. En las referencias (L3 , G2) puede encontrarse un extenso desarrollo de tales técnicas de cálculo.

Desde un punto de vista práctico, es preciso más experiencia computacional que la actualmente disponible para determinar el alcance real de las ventajas e inconvenientes relativos de ambos tipos de métodos. No hay que olvidar que no se conoce la forma analítica de la función dual y por lo tanto, es preciso resolver muchas veces los problemas del primer nivel si se pretende efectuar un proceso de búsqueda

unidimensional en las iteraciones por el método del gradiente en la maximización de la función dual. Por su parte, los métodos de aproximaciones tangenciales requieren la solución de sucesivos programas lineales de crecientes dimensiones. Ello exige disponer de subrutinas de PL y establecer con ellas el necesario transvase de información dentro de la estructura informática de los programas de cálculo.

En cambio, desde un punto de vista de interpretación económica, el uso del método del gradiente en la maximización de la función dual es especialmente interesante.

Supóngase que los distintos subsistemas que componen el problema global representan centros de actividad con ecuaciones características propias y que contribuyen aditivamente al objetivo global. Tales centros usan, en sus respectivos procesos productivos, unos determinados recursos comunes cuyas limitaciones están representadas por las restricciones de interconexión :

$$\sum_{i=1}^S h_i(x_i(t), u_i(t), t) \leq 0 \quad (t = 0, T-1)$$

Bajo tal supuesto, las funciones objetivo de los subproblemas locales del 1^{er} nivel (los Lagrangianos locales), representan la toma en consideración explícita por los centros de actividad del coste de los recursos comunes que deciden emplear a lo largo de sus planes dinámicos de actuación.

Para cada centro i , ($i = 1, S$), este coste está representado por el término :

$$\sum_{t=0}^{T-1} = \sum_{t=0}^{T-1} \left[\sum_{j=1}^{RI} \lambda_j^i(t) h_i^j(x_i(t), u_i(t), t) \right]$$

que se añade a la función objetivo propia de dicho centro.

En su cálculo intervienen :

- 1) las cantidades consumidas por dicho centro i , de cada uno de los bienes j ($j = 1, RI$) sobre los que pesa la restricción de disponibilidad global h_i^j , como consecuencia de los valores que ha decidido dar a sus variables de control $u_i(t)$, $t = 0, T-1$ y los valores correspondientes de sus variables de estado $x_i(t)$, $t = 1, T$, es decir:

$$h_i^j(x_i(t), u_i(t), t) \quad \begin{matrix} t=0, T-1 \\ j=1, RI \end{matrix}$$

- 2) los precios $\lambda_j(t)$, $t = 0, T-1$, $j = 1, RI$ que el centro coordinador del tercer nivel asigna a los bienes sometidos a restricciones comunes.

El proceso de modificación de dichos precios mediante un método del gradiente :

$$\lambda^{0+1}(t) = \lambda^0(t) + \alpha^0 \nabla_{\lambda} \omega(\lambda) \quad (5)$$

es claramente ilustrativo :

- 1) Si la restricción global j está saturada en el instante t :

$$\sum_{i=1}^S h_i^j(x_i(t), u_i(t), t) = 0$$

la suma de las demandas del bien j decididas autónomamente por los centros $i = 1, S$, en el período t , cuando el precio de dicho bien está fijado en $\lambda_j^0(t)$ agota las disponibilidades del mismo y no existe de-

manda insatisfecha. La componente correspondiente del gradiente de la función dual es nula y en consecuencia, el centro coordinador le asignará el mismo precio instantáneo $\lambda_j^{\theta+1}(t)$ dentro del sistema de precios que prepondrá para la iteración siguiente $\theta+1$.

2) Si la demanda no agota sus disponibilidades :

$$\frac{\partial \omega(\lambda)}{\partial \lambda_j^{\theta}(t)} = \sum_{i=1}^S h_i^j(x_i(t), u_i(t), t) < 0$$

por lo que, a través de la ecuación (5), el centro coordinador disminuirá el precio propuesto para la siguiente iteración (sin anularlo necesariamente).

3) Si la demanda supera las disponibilidades, es decir:

$$\frac{\partial \omega(\lambda)}{\partial \lambda_j^{\theta}(t)} = \sum_{i=1}^S h_i^j(x_i(t), u_i(t), t) > 0$$

la relación (5) inducirá un aumento del correspondiente precio.

Los anteriores razonamientos no son sino la interpretación económica, en términos de un mecanismo de ajuste de la demanda por los precios, de un procedimiento numérico de cálculo. Por supuesto, ni el proceso de cálculo ni su interpretación presuponen la existencia y unicidad de las magnitudes $\lambda(k)$, $k = 0, T-1$. Es posible que no exista una trayectoria del vector λ tal que las soluciones independientes de los S problemas de control den conjuntamente la solución del problema global. Dejando para el punto V. 4 el estudio detallado de esta cuestión, analicemos

una última característica de los distintos sistemas de precios generados en los diferentes niveles del proceso.

Los precios de equilibrio de las restricciones de interconexión, $\lambda(k)$, (caso de que existan), son variables duales asociadas a ecuaciones de restricción instantánea que no poseen ningún carácter dinámico puesto que todas las variables que en ellas intervienen se refieren al mismo instante del tiempo. Son fijados y modificados por el centro coordinador sin precisar el conocimiento por su parte ni de los objetivos ni de las ecuaciones que rigen la dinámica propia de cada centro. Finalmente, los sucesivos sistemas de precios generados son puestos en igual conocimiento de todos los centros locales.

Estos, en la solución de sus respectivos problemas, generan unos sistemas de precios de uso interno, puesto que no son transmitidos ni al centro coordinador (que no recibe sino información primal) ni a los demás centros locales. Entre estos figuran especialmente las variables adjuntas resultantes de la solución de cada uno de los problemas de control, asociadas a las variables de estado de cada subsistema. Estos son los únicos precios que tienen carácter dinámico, en el sentido de que las variables que intervienen en las restricciones a las que están asociados, están referidas a distintos instantes del tiempo. También son los únicos que no están sometidos a las restricciones de positividad puesto que tales restricciones (ecuaciones de estado) son del tipo igualdad.

En el apartado siguiente, planteamos la resolución descentralizada del modelo de planificación expuesto en el capítulo IV mediante un algoritmo de este tipo, y aplicamos

en su contexto las anteriores consideraciones. Veremos así, cómo son reproducidos por el proceso numérico de optimización, la coordinación de los centros, la información de que dispone cada uno de ellos y las tareas que le son encomendados en la óptima consecución del objetivo global del sistema.

Analicemos el modelo que fue planteado en el capítulo IV como un problema único de control, para determinar si posee una estructura descomponible que permita la aplicación de los métodos de control jerárquico descritos. Al así hacerlo, será preciso crear nuevas variables de estado y de control para instrumentar la adopción descentralizada de las decisiones que antes se tomaban globalmente por el único y omnipotente centro considerado.

1.- Separabilidad de la función objetivo global.

La función objetivo global propuesta :

$$J = \sum_{t=1}^{T-1} \left[(1+r)^{-t} \sum_{i=1}^N a_i c_i(t)^{b_i} \right] - \alpha \sum_{t=2}^{T-1} \left[l(t) - \sum_{i=1}^N l_i(t) \right] + \sum_{i=1}^N VTK_i [k_i(T) - k_i^*]$$

es aditivamente descomponible en N funciones objetivo, cada una corresponde a uno de los sectores i, (i = 1, N):

$$J_i = \sum_{t=1}^{T-1} (1+r)^{-t} a_i c_i(t)^{b_i} + \alpha \sum_{t=2}^{T-1} l_i(t) + VTK_i [k_i(T) - k_i^*]$$

La cantidad $-\alpha \sum_{t=2}^{T-1} l(t)$, que aparece en la función objetivo global, puede ser eliminada y no aparecer en ninguna de las funciones objetivo sectoriales puesto que es una magnitud exógena cuyo valor es independiente del proceso de optimización.

2.- División del vector de estado y de control en subvectores específicos de cada centro.

jetivo, consideremos a cada sector i , ($i = 1, N$) como un centro autónomo de decisión. Su autoridad responsable, dispondrá del control de unas variables de decisión que le son propias y sometidas a su única autoridad. Estas son :

$$\begin{array}{ll} I_i(t) & (t = 1, T-1) \\ e_i(t) & (t = 2, T-1) \\ BLC_i(t) = m_i(t) - e_i(t) & (t = 1, T-1) \end{array}$$

que en conjunto agotan el vector de control del sistema global.

A su vez, el estado de cada subsistema sectorial está representado por las variables de estado:

$$\begin{array}{ll} K_i(t) & = \text{dotaciones instantáneas de} \\ (t=1, T) & \text{capital} \end{array}$$

$$\begin{array}{ll} \delta_i(t) & = \text{nivel sectorial de deuda ex-} \\ (t=1, T) & \text{terior en el período } t. \end{array}$$

que son específicas del mismo.

Las variables $\delta_i(t)$, ($i=1, N$), no son sino la desagregación en los correspondientes componentes sectoriales de la deuda exterior total del país $\delta(t)$. Estos componentes son generados por las propias trayectorias sectoriales de desarrollo pero no se acumulan, como en el modelo global, en una única cantidad. Evidentemente, sobre ellos pesa la misma restricción limitativa que pesaba sobre su suma: $\delta(t) \leq \delta^*(t)$, lo cual dará lugar a una restricción de interconexión.

Es de la mayor importancia señalar como el análisis descentralizado obliga a considerar más variables de control de las que existen en el modelo global:

En este, los niveles sectoriales de consumo no eran considerados ni variables de estado ni de control. No

eran variables de estado porque no hay una relación dinámica que defina causalmente su valor en $t + 1$ como resultado de la situación del sistema y controles adoptados en t . No eran considerados como variables de control porque al establecer la autoridad central, conjunta y simultáneamente, los niveles de producción, los planes de acumulación de capital y las operaciones exteriores de todos los sectores, los niveles de consumo de estos eran simplemente las variables de holgura que transformaban en igualdad las restricciones de distribución.

Así se obtenían las ecuaciones matriciales de distribución:

$$C(t) = Pq(t) - B I(t) + BLC(t)$$

en donde la autoridad central determinaba simultáneamente:

- a) los niveles de producción de cada sector $q_i(t)$, $(i=1, N)$ a través de las variables de control y de la dotación de capital existente $k_i(t)$ (fruto de anteriores decisiones del tipo b)), y por lo tanto sus exigencias en bienes intermedios.
- b) los recursos destinados a la acumulación de capital en cada sector $\sum_{j=1}^N b_{ij} I_j(t)$, $(i=1, N)$, a través de los valores de las variables de control $I_i(t)$ $(i=1, N)$
- c) las balanzas comerciales no ligadas a la inversión de cada sector $BLC_i(t)$ $(i=1, N)$

Con este procedimiento de determinación simultánea, $C_i(t)$, $(i=1, N)$ era simplemente el resto a cero de las disponibilidades de cada bien i y no precisaba ser conside-

rado como una variable de control suplementaria.

Este enfoque cambia radicalmente cuando se deja a cada centro la libertad de determinar sus propias estrategias de producción y acumulación, con independencia de cuales sean los planes elaborados por los otros sectores.

Para que la autoridad responsable de cada sector, decida los destinos a dar a la producción obtenida, precisa de más variables de control. En efecto, una vez que haya satisfecho sus propias necesidades en:

- a) bienes intermedios por el producidos ($a_{ii} q_i(t)$) precisos para alcanzar el nivel escogido de producción $q_i(t)$
- b) bienes intermedios por el producidos ($b_{ii} I_i(t)$) precisos para desarrollar su programa de inversión
- c) bienes por el producidos destinados al comercio exterior

todavía le queda por hacer una última elección para determinar en qué proporción irá la parte restante a satisfacer su demanda final y en qué proporción será puesta a disposición de la autoridad coordinadora del segundo nivel para satisfacer las necesidades en bienes intermedios de los demás sectores.

Es preciso pues, definir unas nuevas variables de control $C_i(t)$, (niveles sectoriales de consumo), que permitan instrumentar esta última decisión. Estas variables están sometidas a las restricciones individuales

$$C_i^*(t) \leq C_i(t)$$

y participan en las restricciones conjuntas :

$$(1 - a_{ii} - d_{ii}) q_i(t) - b_{ii} I_i(t) + B_i C_i(t) - C_i(t) \geq 0$$

que pueden ser convertidas en igualdades, sin más que dar un nombre ($V_i(t)$) a la variable de holgura correspondiente. Esta significa la parte de producción del sector i que es puesta a disposición de los demás sectores para que estos la utilicen como bien intermedio en sus respectivos procesos de producción y acumulación de capital.

$$(1 - a_{ii} + d_{ii}) q_i(t) - b_{ii} I_i(t) + BLC_i(t) - C_i(t) = V_i(t)$$

Los subvectores de estado y control de cada centro autónomo, que posee autoridad sobre un sector de actividad, son pues :

$$X_i(t) = \begin{bmatrix} k_i(t) \\ \delta_i(t) \end{bmatrix} \quad \mu_i(t) = \begin{bmatrix} I_i(t) \\ \ell_i(t) \\ BLC_i(t) \\ C_i(t) \end{bmatrix}$$

3.- Existencia de ecuaciones de estado específicas de cada subsistema.

La ecuación de estado de cada subsistema sectorial i , que representa su proceso de acumulación de capital:

$$K_i(t+1) = (1 - DPC_i) K_i(t) + \mu_i K_i(t) \left[1 - \left[1 + \frac{\ell_i I_i(t)}{\mu_i K_i(t)} \right]^{-\frac{1}{\ell_i}} \right] \\ (t=1, T-1)$$

depende exclusivamente de variables y parámetros característicos de dicho sector.

Igualmente ocurre con la ecuación de estado que representa el proceso de acumulación de deuda exterior en cada sector i :

$$\delta_i(t+1) = (1 + IDEX) \delta_i(t) + d_{ii} q_i(t) + BLC_i(t) + \Gamma_i I_i(t) \\ (t=1, T-1)$$

y con las funciones de producción :

$$q_i(t) = \eta_i (1 + \nu_i)^t \left[\beta_i [K_i(t)]^{-\rho_i} + (1 - \beta_i) [L_i(t)]^{-\rho_i} \right]^{-\frac{\sigma_i}{\rho_i}}$$

que representan los procesos productivos de cada sector en lo que a inputs primarios de capital y trabajo se refiere.

4.- Sistema de Restricciones específicas de cada subsistema.

Cada sector i está sometido a un conjunto de restricciones instantáneas específicas expresadas exclusivamente en función de sus propias variables :

- Positividad de la inversión : $I_i(t) \geq 0 \quad (t=1, T-1)$

- Positividad de las dotaciones de capital: $K_i(t) \geq 0$
($t=1, T$)

- Limitaciones en el empleo del factor trabajo:

$$0 \leq L_i(t) \leq f_i(K_i(t)) \quad (t=1, T-1)$$

- Limitaciones en el desequilibrio de la balanza exterior de dicho sector por actividades no ligadas a la producción:

$$BLC_i(t) \geq -e_i^*(t) \quad (t=1, T-1)$$

- Restricción de consumo mínimo:

$$C_i(t) \geq C_i^*(t) \quad (t=1, T-1)$$

5.- Restricciones de Interconexión.

Para analizar las restricciones de interconexión, es preciso considerar los recursos limitados cuyo uso es común a todos los sectores y que estos solicitan en cantidades determinadas por sus propios planes de desarrollo. Claramente, estos recursos son de tres tipos:

a) La mano de obra disponible

$$\sum_{i=1}^N l_i(t) \leq l(t) \quad (t=1, T-1)$$

b) La posibilidad de endeudamiento externo:

$$\sum_{i=1}^N \delta_i(t) \leq \delta^*(t) \quad (t=1, T)$$

c) Los bienes producidos por cada uno de los N sectores.

Estos bienes están disponibles en unas cantidades $V_i(t), (i=1, N), (t=1, T-1)$, fijados en cada instante t por cada autoridad sectorial i como consecuencia de los niveles a los que ella misma ha decidido fijar sus actividades de empleo, producción, capitalización, consumo y comercio exterior. Puesto que su sector de origen ya se ha reservado la parte que ha considerado pertinente de su propia producción, las cantidades solamente son demandadas por los sectores $j \neq i, (j = 1, N)$ para atender sus necesidades de bienes intermedios en sus actividades de producción y capitalización.

Las restricciones de interconexión adoptan pues la forma:

$$\sum_{\substack{j=1 \\ j \neq i}}^N a_{ij} q_j(t) + \sum_{\substack{j=1 \\ j \neq i}}^N b_{ij} I_i(t) \leq v_i(t) \quad \begin{matrix} (t=1, T-1) \\ (i=1, N) \end{matrix}$$

Existen pues un total de $N + 2$ restricciones de interconexión instantánea, siendo N el número de sectores considerados en el sistema económico.

6.- Estructura aditiva de las restricciones de interconexión.

Las restricciones de interconexión consideradas poseen la necesaria estructura aditiva. Para explicitarla, particularicemos para el caso que nos ocupa, la formu-

lación general :

$$\sum_{i=1}^N h_i(x_i(t), u_i(t), t) \quad (t=1, T-1)$$

de las restricciones de interconexión utilizada en V. 2.

El número de componentes de h , $h = \sum_{i=1}^N h_i$, es el número de restricciones de interconexión, que allí había sido denominado RI y que aquí es $N + 2$. Por lo tanto, la particularización de la formulación general es:

Para $t = 1, T - 1$:

$$h(x(t), u(t), t) = \begin{bmatrix} h^1(x(t), u(t), t) \\ \vdots \\ h^j(x(t), u(t), t) \\ \vdots \\ h^{RI}(x(t), u(t), t) \end{bmatrix} = \sum_{i=1}^N \begin{bmatrix} h_i^1(x_i(t), u_i(t), t) \\ \vdots \\ h_i^j(x_i(t), u_i(t), t) \\ \vdots \\ h_i^{RI}(x_i(t), u_i(t), t) \end{bmatrix} =$$

$$= \begin{bmatrix} \sum_{i=1}^N h_i^1(x_i(t), u_i(t), t) \\ \vdots \\ \sum_{i=1}^N h_i^j(x_i(t), u_i(t), t) \\ \vdots \\ \sum_{i=1}^N h_i^N(x_i(t), u_i(t), t) \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^N \ell_i(t) \\ \sum_{i=1}^N \delta_i(t) \\ \sum_{j=1, j \neq i}^N a_{ij} q_j(t) + \sum_{j=1, j \neq i}^N b_{ij} I_i(t) \\ \vdots \\ (i=1, N) \end{bmatrix} \leq \begin{bmatrix} \ell(t) \\ \delta(t) \\ V_1(t) \\ \vdots \\ V_N(t) \end{bmatrix}$$

La estructura desarrollada de esta matriz de restricciones de interconexión (figura nºFV.6), muestra como cada una de sus N columnas depende exclusivamente de las variables características del sector al que corresponde.

Se observará también en dicha figura como los elementos diagonales de la submatriz cuadrada ($N \times N$) constituida por las últimas N filas, son nulos. Ello repre-

$\ell_1(t)$	\dots	$\ell_j(t)$	\dots	$\ell_i(t)$	\dots	$\ell_N(t)$	\leftarrow	$\ell(t)$
$\delta_1(t)$	\dots	$\delta_j(t)$	\dots	$\delta_i(t)$	\dots	$\delta_N(t)$	\leq	$\delta(t)$
0	\dots	$a_{ij}q_j(t) + b_{ij}I_j(t)$	\dots	$a_{ji}q_i(t) + b_{ji}I_i(t)$	\dots	$a_{iN}q_N(t) + b_{iN}I_N(t)$	\leftarrow	$V_1(t)$
\vdots	\dots	\vdots	\vdots	\vdots	\vdots	\vdots	\leftarrow	\vdots
$a_{j1}q_1(t) + b_{j1}I_1(t)$	\dots	0	\dots	$a_{ji}q_i(t) + b_{ji}I_i(t)$	\dots	$a_{jN}q_N(t) + b_{jN}I_N(t)$	\leq	$V_j(t)$
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\leq	\vdots
$a_{Li}q_i(t) + b_{Li}I_i(t)$		$a_{Lj}q_j(t) + b_{Lj}I_j(t)$	\dots	0	\vdots	$a_{iN}q_N(t) + b_{iN}I_N(t)$	\leq	$V_i(t)$
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\leftarrow	\vdots
$a_{M1}q_1(t) + b_{M1}I_1(t)$	\vdots	$a_{Mj}q_j(t) + b_{Mj}I_j(t)$	\dots	$a_{Mi}q_i(t) + b_{Mi}I_i(t)$	\dots	0	\leq	$V_N(t)$

FIGURA FV.6 MATRIZ DE LAS RESTRICCIONES DE INTERCONEXION DEL MODELO DE PLANIFICACION.

señala el mecanismo antes explicado por el que el sector i se ha reservado a priori la parte que precisaba de su producción. No concurre pues con su demanda a la saturación de la restricción de interconexión referente a su propio producto.

Estos términos nulos pueden ser substituidos por los segundos términos de las restricciones $V_i(t)$ con signo negativo.

Puesto que la estructura del problema reúne todas las condiciones para ser descompuesta, pasemos a interpretar la aplicación del algoritmo de V. 2.

Definamos un vector de multiplicadores $\lambda(k) \geq 0$, asociado a las $N + 2$ restricciones de interconexión. Denominemos a sus $N + 2$ componentes, de acuerdo con la restricción a la que están asociados, de la siguiente forma:

$\lambda_1(k)$ = variable dual asociada a la restricción = $\lambda_e(k)$
de empleo total de mano de obra en el
instante k

$\lambda_2(k)$ = variable dual asociada a la restricción = $\lambda_f(k)$
de deuda exterior global en el instante
k

$\lambda_{i+2}(k)$ = variable dual asociada a la restricción = $\lambda_{i+2}(k)$
($i=1, N$) disponibilidad del bien i en el instante k.

Con el, podemos construir el Lagrangiano del problema global referido a las únicas restricciones de interconexión:

$$\begin{aligned}
d(x, u, \lambda) = & \sum_{i=1}^N \left[\sum_{t=1}^{T-1} (1+t)^{-t} a_i c_i(t) + \alpha \sum_{t=2}^T u(t) + \gamma \pi_i [k_i(t) - k_i^*] \right] \\
& + \sum_{t=2}^{T-1} \left[\lambda_t(t) \left[\sum_{i=1}^N \ell_i(t) - \ell(t) \right] \right] - \sum_{t=1}^{T-1} \left[\lambda_t(t) \left[\sum_{i=1}^N \ell_i(t) - \delta^*(t) \right] \right] \\
& - \sum_{t=1}^{T-1} \sum_{i=1}^N \left[\lambda_{i+2}(t) \left[\sum_{\substack{j=1 \\ j \neq i}}^N a_{ij} q_j(t) + \sum_{\substack{j=1 \\ j \neq i}}^N b_{ij} I_j(t) - v_i(t) \right] \right]
\end{aligned}$$

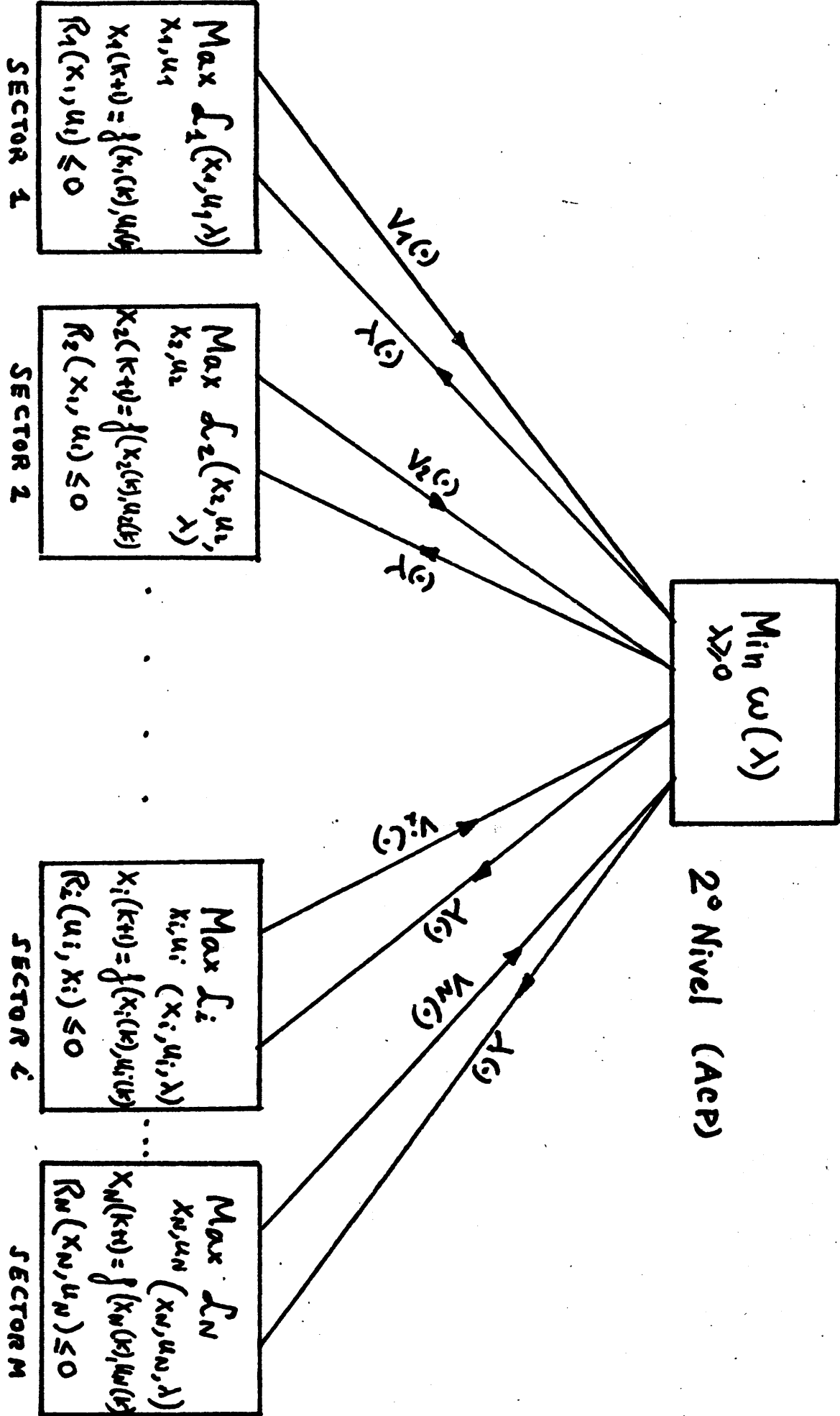
Este Lagrangiano global queda inmediatamente descompuesto en la suma de N Lagrangianos locales, cuya expresión es:

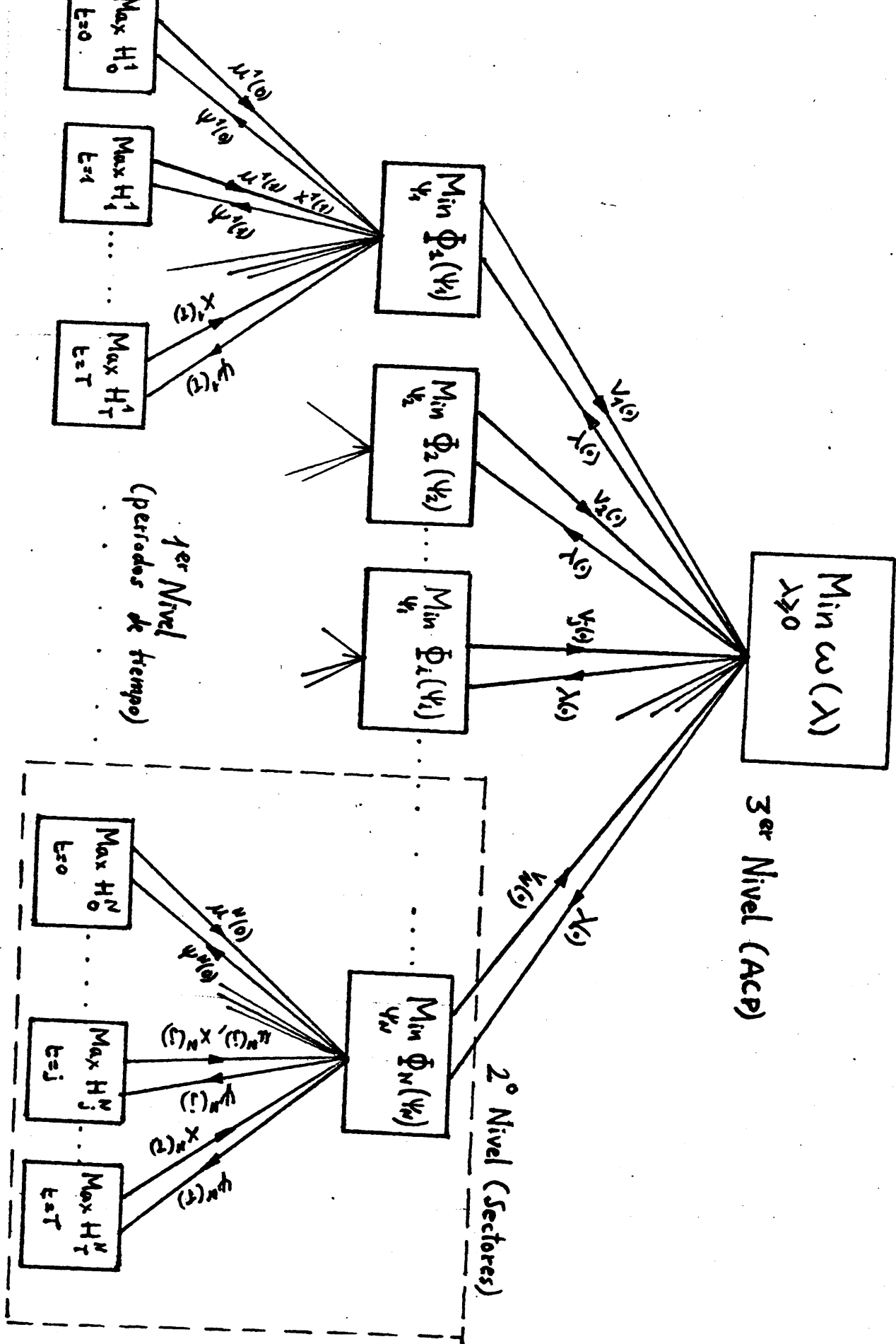
Para $i = 1, N$:

$$\begin{aligned}
d_i(x_i, u_i, \lambda) = & \sum_{t=1}^{T-1} (1+t)^{-t} a_i c_i(t)^{b_i} + \alpha \sum_{t=2}^{T-1} \ell_i(t) + \\
& + \gamma \pi_i [k_i(t) - k_i^*] - \sum_{t=2}^{T-1} \lambda_t(t) \ell_i(t) - \sum_{t=1}^{T-1} \lambda_t(t) \delta_i(t) + \\
& + \sum_{t=1}^{T-1} \lambda_{i+2}(t) v_i(t) - \sum_{t=1}^{T-1} \sum_{\substack{j=1 \\ j \neq i}}^N \left[\lambda_{j+2}(t) [a_{ji} q_j(t) + b_{ji} I_j(t)] \right]
\end{aligned} \tag{10}$$

Estos Lagrangianos son las funciones objetivo de cada uno de los centros locales en los que se descompone el primer nivel en la estructura jerárquica a dos niveles representada en la figura nº FV.4

Como en ella se observa, el primer nivel está constituido por tantos centros de decisión independientes como sectores se hayan considerado en la desagregación de la economía. La autoridad responsable de cada sector, construye su propia estrategia de desarrollo mediante la resolución de un problema de control óptimo, absolutamente independiente de los otros $N - 1$ problemas del mismo tipo que están resolviendo si-





multáneamente los demás centros.

La funcional objetivo de cada uno de los centros del primer nivel es el Lagrangiano local correspondiente dado por la fórmula (10). Las ecuaciones de estado, variables y restricciones de cada uno de estos problemas de control, son específicos del mismo.

Jerárquicamente situado por encima de los centros locales, aparece el centro coordinador del segundo nivel que cumple las funciones de una "Autoridad Central del Plan". De ella depende la gestión de las restricciones comunes a todos los centros, que estos ignoran en sus procesos locales. Su función básica consiste en la modificación del vector $\lambda(\cdot)$ de forma que se consiga la satisfacción de tales restricciones de interconexión. El proceso de modificación lo efectuará de forma a minimizar la función dual $\omega(\lambda)$:

$$\omega(\lambda) = \underset{x, u}{\text{Max}} \mathcal{L}(x, u, \lambda)$$

(observese que, a diferencia del planteamiento de V. 2, el objetivo primal es aquí de maximizar la función objetivo).

La algorítmica de cálculo es idéntica a la presentada en dicho apartado. Sin repetir aquí los pasos del proceso numérico, demosle una interpretación económica en el contexto del modelo de planificación planteado. Dividimos la exposición en los siguientes puntos :

1) Precios de bienes producidos y precios de bienes de capital.

Existe una serie de importantes diferencias entre ellos, debidas a la forma en que son generados en el modelo. Los precios de los bienes producidos por los distintos

sectores son fijados por la Autoridad central del plan (centro coordinador del segundo nivel). Son los multiplicadores asociados con las restricciones de disponibilidad común de dichos bienes, $\lambda_{i+2}(\cdot)$, $(i=1, N)$. Pero la Autoridad central (ACP) también fija un precio a dos bienes de uso común que no son producidos específicamente por ningún sector: la fuerza de trabajo y las posibilidades de endeudamiento externo. Sus precios son respectivamente $\lambda_1(\cdot)$ y $\lambda_2(\cdot)$. Al fijar $\lambda_1(\cdot)$ y $\lambda_2(\cdot)$ la ACP está fijando el tipo de salario y el tipo de cambio.

Este sistema de precios constituye el único tipo de información que la ACP transmite a los centros locales. Estos reciben pues exclusivamente información dual. Tal sistema no posee una intrínseca naturaleza dinámica, (aunque sí son variables en el tiempo), porque las restricciones que los generan son restricciones instantáneas en las que todas las variables están referidas al mismo instante del tiempo.

Por el contrario, los precios de los bienes de capital son fijados por las autoridades locales de cada sector como resultado de la solución de su propio problema de control óptimo. Tienen una intrínseca naturaleza dinámica puesto que son las variables adjuntas asociadas a las ecuaciones de estado que describen el proceso de acumulación de capital en cada sector. No son conocidos de la ACP ni de ninguno de los demás centros de decisión sectoriales (CDS).

- 2) Información necesaria a e información facilitada por cada centro.

de control óptimo, los CDS reciben exclusivamente información dual por parte de la ACP. Concretamente, esta les notifica el precio que atribuye al bien que el sector produce y los precios que asigna a los bienes que el CDS precisa para llevar a cabo sus planes (trabajo, reservas de cambio y bienes intermedios procedentes de los demás sectores). Recuérdese que los sectores son centros específicos de producción y no de consumo, por lo que sólo precisan bienes intermedios destinados a la producción y acumulación pero no demandan bienes procedentes de los demás sectores para satisfacer demandas finales. Por supuesto, cada CDS conoce sus ecuaciones de estado, funciones de producción, objetivo y restricciones propias.

En cambio, la ACP precisa disponer de una información tecnológica mínima y solamente recibe una muy limitada información de tipo primal por parte de los CDS.

En efecto, los únicos datos que precisa conocer son las matrices de input-output (matriz A) y de coeficientes de capital (matriz B) de la economía. Más precisamente, tampoco precisa conocer los elementos diagonales de ambas matrices. La información que recibe de los centros se limita a dos valores: sus niveles totales de producción $q_i(t)$ en cada instante y los excedentes $v_i(t)$ que, después de satisfacer todas sus necesidades, en la medida por ellos mismos determinada, van a poner a disposición de los demás sectores, al precio $\lambda_i(t)$ fijado por la ACP.

3) Simulación de las transferencias entre ACP y CDS.

Analizando las funciones objetivo que los CDS (los Lagrangianos locales) se proponen maximizar con respecto a sus propias variables, dado $\lambda(t)$, y la función dual que

la ACP debe minimizar con respecto a su sistema de -
precios $\lambda(\cdot)$, dadas las propuestas primales, es po-
sible reconocer en los componentes de $\lambda(\cdot)$ los típi-
cos precios de cesión o transferencia.

En efecto, los Lagrangianos locales no son sino -
el resultado de añadir a la verdadera función objeti-
vo de cada sector, el balance entre el importe de los
bienes por él cedidos a la ACP y el coste de los fac-
tores de producción que el sector adquiere a la ACP
y que esta ha "comprado" a los demás sectores (bie-
nes intermedios) o que gestiona centralizadamente (re-
servas de cambio y fuerza de trabajo). Los CDS fijan
las cantidades físicas en las que se sitúan los nive-
les de tales intercambios y la ACP fija los precios
unitarios del mismo. Tales elementos están represen-
tados por los siguientes términos de la expresión -
(10):

$$-\sum_{t=2}^{T-1} \lambda_i(t) l_i(t)$$

Coste para el CDS nº i -
del factor trabajo a lo
largo de los T-2 perio-
dos del plan (recuérdese
que la situación inicial
de empleo viene dada).

$$-\sum_{t=1}^{T-1} \lambda_r(t) r_i(t)$$

Id. Id. de la disponibi-
lidad de reservas de cam-
bio en los T-1 periodos -
del plan.

$$-\sum_{t=1} \sum_{\substack{j=1 \\ j \neq i}} \left[\lambda_{j+2}(t) [a_{ji} q_i(t) + b_{ji} I_i(t)] \right]$$

Id. Id. de los bienes intermedios procedentes de los demás sectores $j, (j=1, N)$ ($j \neq i$)

(Cada bien j es -
facturado a su -
precio instantáneo

$$\lambda_{j+2}(t)$$

decidido por la ACP.)

$$\sum_{t=1}^{T-1} \lambda_{i+2}(t) V_i(t)$$

Importe que la ACP abona -
al CDS no i por la cesión -
de $V_i(t)$ unidades de los
bienes producidos por este.

Es decir, cada CDS trata de maximizar su objetivo incluyendo en este el balance de sus transferencias con la ACP.

Esta, a su vez trata de minimizar la función dual que no representa sino el balance de sus transferencias con todos los centros.

Si la ACP utiliza un método del gradiente en la resolución del problema en su segundo nivel, la variación del sistema de precios que el algoritmo induce:

$$\lambda^{0+1} = \lambda^0 - \alpha \nabla_{\lambda} \omega(\lambda)$$

tiende claramente a conseguir el equilibrio de las -
 restricciones de interconexión, es decir el equili-
 brio en el mercado de los bienes de uso común por los
 CDS. En efecto, recuerdese que, bajo la hipótesis de
 diferenciabilidad de la función dual, sus derivadas
 parciales con respecto a los elementos $\lambda_i(k)$
 son los grados de insatisfacción de las restriccio-
 nes de interconexión. El precio de un factor no será
 modificado de iteración en iteración sino cuando el
 mercado de dicho bien este en equilibrio, la restric-
 ción esté saturada y la correspondiente derivada par-
 cial sea nula.

Todas las restricciones del problema se verán satisfe-
 chas en el optimo, (si este llega a alcanzarse). En rea-
 lidad, las restricciones locales de los sub-problemas -
 de control en los que se descompone el primer nivel, pue-
 den verse satisfechas a lo largo de todas las iteracio-
 nes si en su resolución se utiliza un método primal. -
 Cualquiera que sea el método empleado en las soluciones
 de los problemas locales, las propuestas que los CDS de-
 ben de presentar a la ACP deben ser sus optimos globales
 correspondientes a cada propuesta de precios por parte
 de esta. Evidentemente ello plantea problemas de viabi-
 lidad del método cuando los sub-problemas locales no son
 programas convexos.

Guardando este tema para el siguiente apartado, es -
 evidente que la convergencia se verá facilitada cuanto -
 más cerca esté la solución inicial de la óptima. Normal-
 mente, se toma $\lambda(k) = 0$ como trayectoria inicial del -
 sistema de precios de la ACP. Pero si se conoce cuales

son las restricciones de interconexión que serán normalmente activas, se puede definir un sistema inicial de precios más acorde con las características del problema. Por ejemplo, en el modelo planteado, sabemos que será difícil para el sistema económico absorber el desempleo que padece, al menos en los primeros periodos del plan. En consecuencia, es razonable tomar $\lambda_i(\cdot) = 0$. En cambio, los bienes producidos por el sector industria pesada (I.P.) serán sin duda limitativos en la elaboración de estrategias de desarrollo. De acuerdo con ello, el sistema inicial debe acordar un valor no nulo al precio de dichos bienes $\lambda_{IP}(\cdot)$. Igualmente debe ocurrir con el tipo de cambio $\lambda_1(\cdot)$.

Como hemos indicado en la introducción a este capítulo, el procedimiento de resolución por descomposición jerárquica que hemos expuesto, simula un comportamiento completamente descentralizado del sistema económico puesto que la autoridad del primer nivel no genera propuestas primales. Estas son el resultado de las decisiones de los centros de autoridad local. También hemos indicado en la introducción, que, desgraciada y previsiblemente, tal método no es capaz de resolver todos los problemas. En otras palabras, la estructura descomponible de un problema no garantiza que su solución pueda obtenerse mediante su descomposición jerárquica bajo un centro coordinador que genere solamente información dual.

El punto básico es la existencia o no de un punto de silla del Lagrangiano del problema global. Tal existencia puede garantizarse si el problema global es un programa convexo. El punto siguiente lo dedicamos al análisis de -

las condiciones de aplicabilidad de estos métodos, más -
alla de las condiciones de suficiencia representadas por
la convexidad.

Exponiendo el concepto de "gap" y utilizando los teo-
remas de Everett, analizamos la importancia de la informa-
ción generada por tales métodos, aún en el caso de que no
obtenham la solución del problema específicamente propues-
to.

V. 4 POSIBILIDADES, LIMITES E INTERES DE LA APLICACION DEL METODO.

El método de descomposición jerárquica de problemas de
control óptimo que hemos considerado, pretende construir
la solución descentralizada del problema global mediante
el mecanismo coordinador de un sistema de precios. El al-
goritmo que lo instrumenta, es a su vez un procedimiento
de cálculo del valor óptimo de dichos precios, a través -
de la maximización de la función dual por el centro coor-
dinador del segundo nivel.

Tal proceso de cálculo parte del supuesto de la exis-
tencia de este sistema de precios. Pero, ¿ puede esta ser
asegurada a priori para todos los problemas que se preten-
da resolver ?.

Precisamente porque sabemos que la respuesta es no, es
preciso, desde un punto de vista práctico, desglosar la -
pregunta en otras dos:

que las soluciones, óptimas localmente, propuestas por los subsistemas:

$$x_1^*(\lambda(\cdot)), x_2^*(\lambda(\cdot)), \dots, x_N^*(\lambda(\cdot))$$

bajo los condicionamientos por el representados, constituyan una solución del problema global?

b) ¿ Como puede reconocerse la optimalidad de un sistema de precios generado por el algoritmo?

La primera de dichas preguntas no tiene una respuesta universalmente válida. La existencia solamente puede ser - garantizada a priori para los problemas que constituyan - programas convexos. Pero pueden también existir en programas no convexos.

Observese que el proceso iterativo a dos niveles, (minimización descompuesta del Lagrangiano global con respecto a las variables primales + maximización de la función dual con respecto a las variables duales), está diseñado, no para acceder directamente al óptimo del problema, sino para localizar el posible punto de silla de Lagrangiano - global. Puesto que el componente primal x^* , del punto de silla (x^*, λ^*) del Lagrangiano de un programa matemático que lo posea, es una solución del mismo (teorema de suficiencia del punto de silla (-L5 -)), la localización del punto de silla equivale a la resolución del problema. Pero puesto que la condición es suficiente y no necesaria, tal punto de silla puede no existir. Es evidente que, en - tal caso, un algoritmo de este tipo no puede facilitar la solución.

La existencia de puntos de silla estando garantizada - para programas convexos, los métodos expuestos son directamente aplicables a este tipo de problemas. El recíproco no es cierto; en problemas no convexos no está garantizada la imposibilidad de su aplicación.

En estas condiciones, la respuesta a la segunda pregunta reviste extraordinario interés y es, afortunadamente, independiente del problema aplicado.

En efecto, es bien sabido y no es preciso detenerse en demostrarlo (L5, L12) que si $\lambda^*(\cdot)$ es el sistema de precios para el que se alcanza el máximo de la función - dual y $x^*(\cdot)$ es una propuesta primal correspondiente \dagger , esta es la solución óptima del problema global si y solo si los valores óptimos de las funciones objetivos primales y duales son iguales.

Dotados de un test fiable y de fácil aplicación para - testar la optimalidad de los resultados, se plantean dos tipos de cuestiones:

- a) analizar teóricamente las condiciones, otras que las de convexidad, de aplicabilidad del método.
- b) analizar la utilidad de la información generada en el caso de que las soluciones hacia las que se converja - (en el caso en que la convergencia se alcance), no se revelen como óptimas.

\dagger Dependiendo de la naturaleza de los problemas locales, nada garantiza la unicidad de sus propuestas para un - mismo sistema de precios

La respuesta a la primera cuestión es debida a los trabajos de Luenberger (-L12-) y Rockafellar (-R5-) y constituye una interesante profundización en el significado de la dualidad. Sin embargo, no conduce a ningún test a priori sobre la aplicabilidad del método.

Volviendo más tarde a este tema, la segunda cuestión es especialmente interesante en las aplicaciones económicas.-- Los razonamientos que siguen se basan en la siguiente consideración básica que ya hemos enunciado en anteriores capítulos: los valores de los parámetros de modelos económicos, y los datos que los alimentan se conocen con un elevado grado de imprecisión. En consecuencia, es preciso admitir que:

- a) el problema específicamente planteado debe considerarse como un elemento representativo de una familia de problemas dotados de similar estructura, obtenidos al variar los valores de sus parámetros.
- b) más que en la solución del problema para un conjunto concreto de valores de sus parámetros, nuestro interés se centra en el conocimiento de la sensibilidad de las soluciones de los elementos de la familia frente a variaciones en tales valores.

Aceptando este enfoque, que parece impuesto por una apreciación realista del estado en el que se encuentra actualmente la construcción de modelos económicos, la aplicación de los métodos duales pasa a cobrar un nuevo interés.



Para explicar el mismo, es preciso recordar los teoremas enunciados por Everett (^{E2} -). En su presentación, - modificada para tener en cuenta la descomposición del problema del primer nivel, nos referimos a la más simplificada forma de presentar el problema de control como un programa matemático:

$$\text{Min}_x f(x) = \text{Min}_x \sum_{i=1}^N f_i(x_i)$$

$$g(x) = \sum_{i=1}^N g_i(x_i) \leq 0$$

$$x = (x_1 : x_2 : x_3 : \dots : x_N) \in S = S_1 \times S_2 \times \dots \times S_N$$

donde:

x vector primal del problema global que contiene a las variables de estado y control de cada centro.

x_i vector de variables de estado y control específico - de cada centro i
($i=1, N$)

$g(x)$ vector de restricciones comunes aditivas, de dimensión m

S_i conjunto de soluciones factibles de cada problema - local de control, ($i=1, N$), definido por las ecuaciones de estado y las restricciones propias de cada uno de ellos.

S producto cartesiano de los S_i conjuntos locales. Representa el conjunto de soluciones primales factibles cuando no se consideran las restricciones de interconexión.

Si el conjunto de propuestas primales $x_i^*(\lambda), (i=1, N), \lambda \geq 0$ resuelven el problema Lagrangiano del primer nivel:

$$\left. \begin{array}{l} \min_{x_i} f_i(x_i, \lambda) \\ x_i \in S_i \end{array} \right\}$$

estas propuestas constituyen la solución del siguiente - problema global :

$$\min_x f(x) = \min_x \sum_{i=1}^N f_i(x_i)$$

$$g^j(x) = \sum_{i=1}^N g_i^j(x_i) \leq y_j \quad \begin{array}{l} (j=1, m) \\ (i=1, N) \end{array}$$

en donde:

$$y_j = g^j(x^*(\lambda)) = \sum_{i=1}^N g_i^j(x_i^*(\lambda)) \quad \text{si } \lambda_j > 0$$

$$y_j \geq g^j(x^*(\lambda)) = \sum_{i=1}^N g_i^j(x_i^*(\lambda)) \quad \text{si } \lambda_j = 0$$

Observe que este problema unicamente difiere del originalmente planteado, en los segundos miembros de las restricciones de interconexión

Puesto que las restricciones de interconexión las hemos interpretado, en el contexto del modelo de planificación - propuesto, como las restricciones en la disponibilidad de bienes comunes a todos los sectores, y las variables duales $\lambda_j(0)$ como los precios de los bienes cuyas limitaciones

traducen las restricciones g^j correspondientes, el problema que resuelve la propuesta de los CDS, para un sistema dado de precios $\lambda(\cdot) \geq 0$ es uno en el que:

- 1) Las cantidades disponibles de los bienes j a los que se asigna, a través del sistema $\lambda(\cdot)$ un precio > 0 son iguales a las cantidades $\sum_{i=1}^n g_i^j(x_i^*(\lambda))$ demandadas por las propuestas primales $x_i^*(\lambda)$. Estas demandas son pues las cantidades máximas que permite consumir el problema modificado.
- 2) Las cantidades disponibles de los bienes j a los que el sistema λ asigna un precio nulo, son por lo menos iguales (pero pueden ser mayores) a las cantidades $\sum_{i=1}^n g_i^j(x_i^*(\lambda))$ demandadas por las propuestas primales $x_i^*(\lambda)$. Estas demandas representan pues las disponibilidades mínimas con las que se plantea el problema modificado. El interés del teorema es el siguiente:

Cada vez que el centro coordinador del segundo nivel (ACP) genera un sistema de precios y lo transmite a los CDS, estos resuelven descentralizadamente el problema Lagrangiano y al así hacerlo, encuentran la solución óptima de un problema perteneciente a la familia del propuesto, cuyos elementos componentes difieren entre sí en el valor del segundo miembro de las restricciones de interconexión.

Así, en la búsqueda de la solución de un problema específico, obtenemos, inevitablemente y sin coste suplementario de cálculo, las soluciones de tantos problemas de la familia del propuesto como iteraciones

se efectúen entre los dos niveles del algoritmo.

Cuando no se conoce con certeza el segundo miembro de las restricciones conjuntas, es de la mayor importancia para un análisis de sensibilidad, el conocimiento de las soluciones correspondientes a toda una serie de problemas que difieran solamente en los valores - de estos términos.

La utilidad de esta información depende de la diferencia que exista entre los segundos miembros del problema original y los generados en correspondencia a las soluciones obtenidas en las sucesivas iteraciones. A partir de una propuesta dual $\lambda(\cdot)$ se puede generar un sistema de segundos miembros muy alejados de los originales, pero el 2º teorema de Everett nos informa de las modificaciones que es preciso efectuar en tal sistema de precios para que el segundo miembro del problema modificado resultante (del que necesariamente obtendremos la solución) se aproxime al segundo miembro del problema original.

Teorema 2 de Everett

Dados dos sistemas de precios generados por la ACP tales que:

$$\lambda^1(\cdot) \geq 0, \lambda^2(\cdot) \geq 0$$

$$\lambda_j^2 > \lambda_j^1$$

$$\lambda_i^2 = \lambda_i^1 \quad \forall i = 1, m, i \neq j$$

si $x(\lambda^1), x(\lambda^2)$ son las correspondientes propuestas primales generadas por el primer nivel como resultado de la resolución descentralizada del problema Lagrangiano, se verifica:

$$g^j(x(\lambda^2)) = \sum_{i=1}^N g_i^j(x_i(\lambda^2)) \leq g^j(x(\lambda^1)) = \sum_{i=1}^N g_i^j(x_i(\lambda^1))$$

El teorema asegura la certeza de la intuitiva hipótesis de que, al aumentar el precio asignado a un factor, mantenién dose constantes los demás, el uso total del mismo - propuesto por el conjunto de los centros, tenderá a disminuir.

Así podemos generar segundos miembros del problema modificado más próximos a los del problema original. Evidentemente, no todos los vectores del segundo miembro de las restricciones de interconexión pueden generarse por este - procedimiento, de lo contrario estaría asegurada la solución del mismo problema original. En la terminología anglosajona (- ^{E2} -) se denominan "gaps" a aquellos vectores segundo - miembro, o conjunto de los mismos, para los que no se puede resolver su correspondiente problema asociado.

Por otra parte, la función dual posee propiedades que - permiten interpretar de forma ilustrativa la secuencia de sus valores obtenidos en el proceso de cálculo.

En efecto, si bien la igualdad de los valores de las - funciones objetivo primales y duales solamente se verifica para las soluciones óptimas, existe un conocido teorema - (- ^{L5} -) que garantiza la desigualdad:

$$\omega(\lambda) \leq f(x) \quad \forall \lambda \geq 0, x \text{ factible}$$

Es decir, el valor de la función dual es una cota inferior del valor óptimo del problema primal, para cualquier sistema de precios de coordinación $\lambda(\cdot) \geq 0$ y cualquier solución primal factible.

Ello implica que el procedimiento de cálculo genera una secuencia de cotas superior e inferior del valor óptimo de la función objetivo del problema global a partir del momento en que se generen propuestas de los CDS que sean factibles globalmente:

$$\omega(\lambda) \leq f(x^*) \leq f(x)$$

x^* solución primal del problema global

$$\forall \lambda \geq 0$$

$$\forall x \left\{ \begin{array}{l} x_i \in S_i \\ \sum_{i=1}^N g_i(x_i) \leq 0 \end{array} \right.$$

Desde un punto de vista de práctica numérica de solución, este resultado es de fundamental importancia. En efecto, en la práctica solamente se obtienen soluciones aproximadas. En consecuencia, es necesario un procedimiento que indique cuál es el grado de aproximación obtenido en la solución y que pueda utilizarse como test determinante del final del proceso iterativo cuando la solución actual se sitúe dentro de cierto entorno del óptimo.

Ningun método primal genera este tipo de información en el curso de su aplicación. Por ello, y dadas las condiciones en las que se plantean los modelos económicos, creemos que los métodos duales de descomposición jerárquica son del mayor interés aunque no pueda asegurarse a priori su eficacia en la resolución de un problema concreto.

Aquí entroncamos de nuevo con la cuestión a) que nos - habíamos planteado anteriormente en este mismo apartado. Como hemos indicado, Luenberger y Rockafeller han caracterizado las condiciones que deben reunir problemas no convexos para que el método pueda encontrar las soluciones de las mismas.

Dado el interés de tales resultados y la poca difusión de la que han gozado en la literatura económica, pasamos a exponerlos brevemente y a ilustrarlos gráficamente. Tal exposición está tomada fundamentalmente de la que presentan Luenberger (-L12-), Héstenes (-H3-) y Lasdon (-L3-). Las figuras presentadas, quizá las más ampliamente reproducidas en la literatura de los métodos avanzados de programación matemática, son las que aparecen en las citadas referencias.

El punto fundamental es la existencia de hiperplanos soportes a un determinado conjunto C que pasamos a definir:

Dado el problema primal:

$$\begin{array}{ll} \text{Min} & f(x) \\ x & (x \in E^n) \\ g(x) & \leq 0 \\ x \in S & (g(x) \text{ es un vector de dimensión } m) \end{array}$$

considerémoslo incluido en una familia de problemas con segundos miembros de las restricciones $g(x)$ parametrizado por el vector b

$$\begin{array}{l} \text{Min} & f(x) \\ x & \\ g(x) & \leq b \\ x \in S & \end{array}$$

El problema primal propuesto es el elemento de dicha familia correspondiente a $b=0$

Definamos la función:

$$\Phi(b) = \min_x \{ f(x) \mid g(x) \leq b, x \in S \}$$

Cuyo dominio de definición es aquel conjunto F de vectores b para los que el problema primal correspondiente - tiene solución:

$$F = \{ b \mid \exists x \in S, g(x) \leq b \}$$

Evidentemente, $\Phi(b)$ es una función no creciente del vector b , cualquiera que sean las condiciones de convexidad del problema primal.

El conjunto $C \in E^{m+1}$ es el conjunto de puntos (b_0, b) delimitado inferiormente por $\Phi(b)$:

$$C = \{ (b_0, b) \mid b \in F, b_0 \geq \Phi(b) \}$$

Si las funciones f , g y el conjunto S son convexos, se demuestra que los conjuntos F , C y la función Φ también lo son. En tal caso la función Φ y el conjunto C pueden ser ilustrados mediante una representación gráfica como la de la figura FV.7.

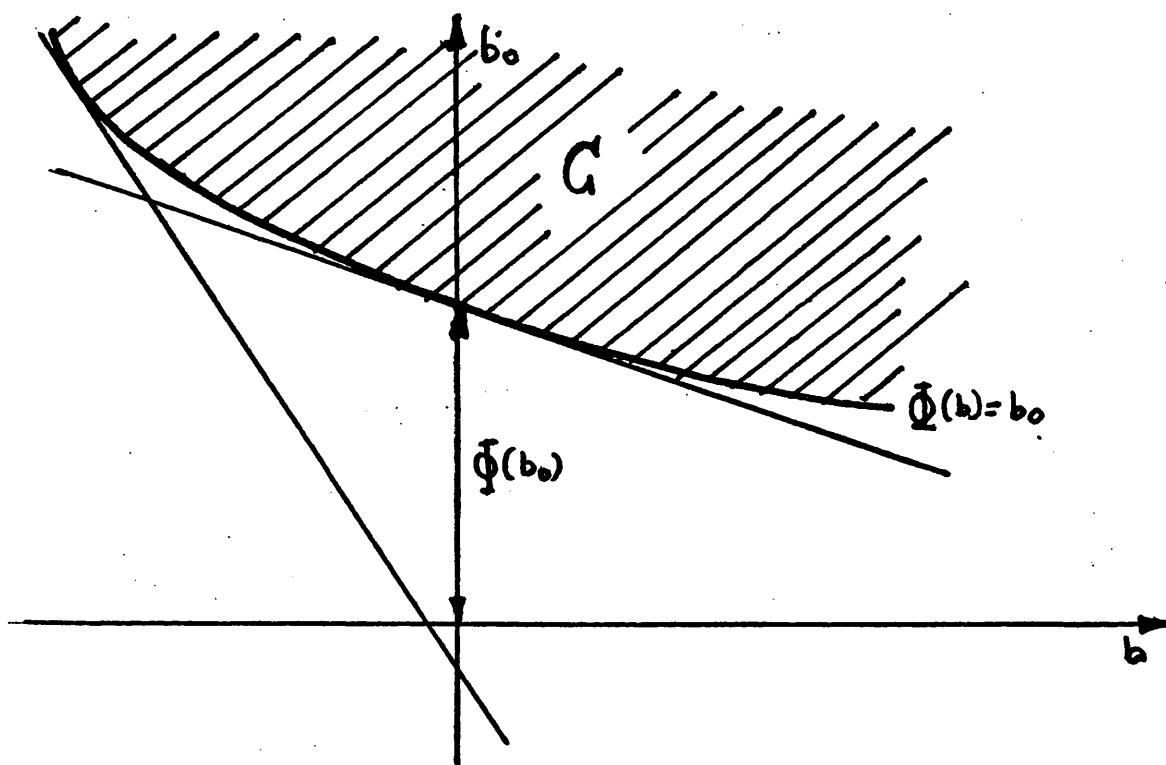
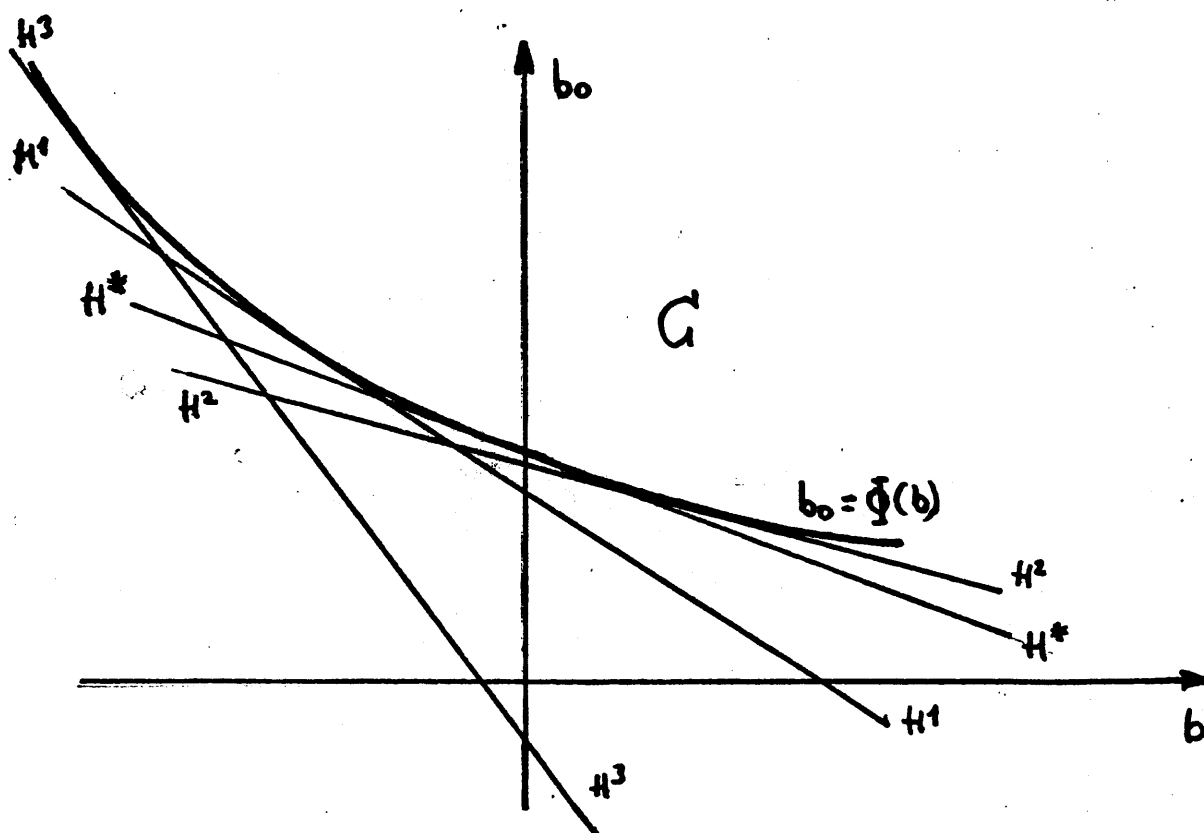


FIGURA FV.8 ILUSTRACION DEL TEOREMA DEL HIPERPLANO MINIMIZANTE



En ella se pone de manifiesto la relación existente entre métodos duales y los hiperplanos soportes del conjunto C, - mediante el siguiente teorema, cuya demostración puede encontrarse en (L 1 2, L 3 -) y cuyo enunciado está modificado para adaptarlo a nuestros propósitos:

Teorema del Hiperplano minimizante:

Sean x_i^* ($i=1, N$) un conjunto de soluciones factibles de los N problemas del primer nivel, es decir:

$$x_i^* \in S_i \quad (i=1, N)$$

$$x^* = [x_1^* : x_2^* : x_3^* : \dots : x_N^*] \in S_1 \times S_2 \times \dots \times S_N = S$$

Sea $\lambda^* \geq 0$ un sistema de precios propuesto por la - autoridad coordinadora del segundo nivel.

El conjunto de propuestas x_i^* ($i=1, N$) constituyen una solución al problema del Lagrangiano global, es decir:

$$\mathcal{L}(x^*, \lambda^*) = \sum_{i=1}^N \mathcal{L}_i(x_i^*, \lambda^*) = \min_{x \in S} \mathcal{L}(x, \lambda^*) = \min_{x_i \in S_i} \left[\sum_{i=1}^N \mathcal{L}_i(x_i, \lambda^*) \right]$$

si y solamente si:

el conjunto de puntos del espacio E^{m+1} :

$$H = \{ (b_0, b) \mid b_0 + \lambda^* b = \mathcal{L}(x^*, \lambda^*) \}$$

es un hiperplano soporte del conjunto C en el punto $(f(x^*), g(x^*))$, es decir, el punto de coordenadas:

$b_0 = f(x^*)$: valor de la función objetivo primal cuando el vector primal es el conjunto de propuestas locales.

$$b = g(x^*)$$

vector del segundo miembro de las -
restricciones de interconexión que
indica una disponibilidades máximas
de recursos iguales a las demandas
representadas por las propuestas

De acuerdo con este resultado, la solución de los tres
tipos de problemas que estamos considerando simultaneamente
poseen la siguiente interpretación geométrica, ilustrada -
en la figura FV.8

a) Problema Lagrangiano $\equiv \min_{x \in S} \mathcal{L}(x, \lambda^*) \quad \lambda^* \geq 0$

Consiste en encontrar un número real $\omega(\lambda^*)$:

$$\omega(\lambda^*) = \min_{x \in S} \mathcal{L}(x, \lambda^*) = \mathcal{L}(x^*, \lambda^*)$$

tal que el hiperplano cuya pendiente está dada por el -
vector $-\lambda^*$ y cuyo ordenada en el origen es $\omega(\lambda^*)$,
sea un hiperplano soporte del conjunto C

b) Problema dual $\equiv \max_{\lambda \geq 0} \omega(\lambda)$

Encontrar el hiperplano soporte con pendiente $-\lambda \leq 0$
que tenga la máxima ordenada en el origen

c) Problema primal $\equiv \min_{x \in S} f(x)$
 $g(x) \leq 0$

Encontrar el punto $(b_0, 0)$ del conjunto C con -
menor ordenada en el origen, es decir :

$$b_0 = \Phi(0) = \min_x \{ f(x) \mid x \in S, g(x) \leq 0 \}$$

Puesto que la función $\Phi(b)$ es no creciente en b , todos los hiperplanos soportes al conjunto C tienen pendientes negativas y, como explicita la figura Fv.3, la ordenada en el origen de todos ellos es menor o igual que el valor óptimo del problema primal propuesto $\Phi(0)$. Lo cual no es sino la traducción geométrica de la propiedad antes utilizada de la función dual para acotar la solución del problema:

$$\omega(\lambda) \leq \min f(x)$$

La igualdad entre $\max_{\lambda \geq 0} \omega(\lambda)$ y $\min \{f(x) \mid x \in S, g(x) \leq 0\}$ solamente podrá verificarse cuando el conjunto C posea un hiperplano soporte en el punto

$$(b_0, 0) = (\Phi(0), 0)$$

cuya ordenada en el origen, la máxima de todos los hiperplanos soportes al conjunto C , $(\max_{\lambda \geq 0} \omega(\lambda))$, y evidentemente igual a $\Phi(0)$, iguale el valor óptimo del primal:

$$\Phi(0) = \min_x \{f(x) \mid x \in S, g(x) \leq 0\}$$

La existencia del hiperplano soporte en el punto $(\Phi(0), 0)$, como en cualquier otro punto $(\Phi(b), b)$ de la frontera del conjunto C , está garantizada por la convexidad del mismo derivada a su vez de la convexidad del problema primal.

Si este no es convexo, tampoco lo es la función $\Phi(b)$ ni el conjunto C . Pero, dependiendo de su forma, puede existir o no un hiperplano soporte en el punto de interés $(\Phi(0), 0)$. La solución podrá obtenerse si este existe, como en el caso

representado por la figura FV.9 ; y no podrá serlo si el conjunto C no lo posee como en el caso de la figura FV.10

En general, el método de descomposición dual podrá resolverse para todos los elementos de la familia de problemas cuyos puntos representativos en el gráfico de la función $\Phi(b)$ se situen en la intersección de esta con la envoltura convexa del conjunto C. Aquellos para los que esto no ocurra constituyen los "gaps" antes citados. Es decir, los vectores del segundo miembro de las restricciones de interconexión que no pueden ser generados en el transcurso de las iteraciones entre los dos niveles de la jerarquía.

Puede ser, justamente, argumentado que tal caracterización de los problemas a los que puede ser aplicado el método es una elegante construcción teórica sin demasiada relevancia práctica. En efecto, la existencia de un hiperplano soporte del conjunto C en el punto solución no puede ser testada a priori de una manera práctica y general, luego no es posible derivar de la teoría expuesta, un test a priori de la aplicabilidad del método.

Este es el momento preciso para recordar aquello que el uso habitual del término parece haber conseguido hacer olvidar: tampoco existe un procedimiento fácilmente aplicable para testar la convexidad de una función no lineal con un número elevado de variables. Por lo tanto, definir la aplicabilidad de los métodos duales sobre las condiciones de convexidad del problema es poco más operativo que definirla sobre la existencia de hiperplanos soportes.

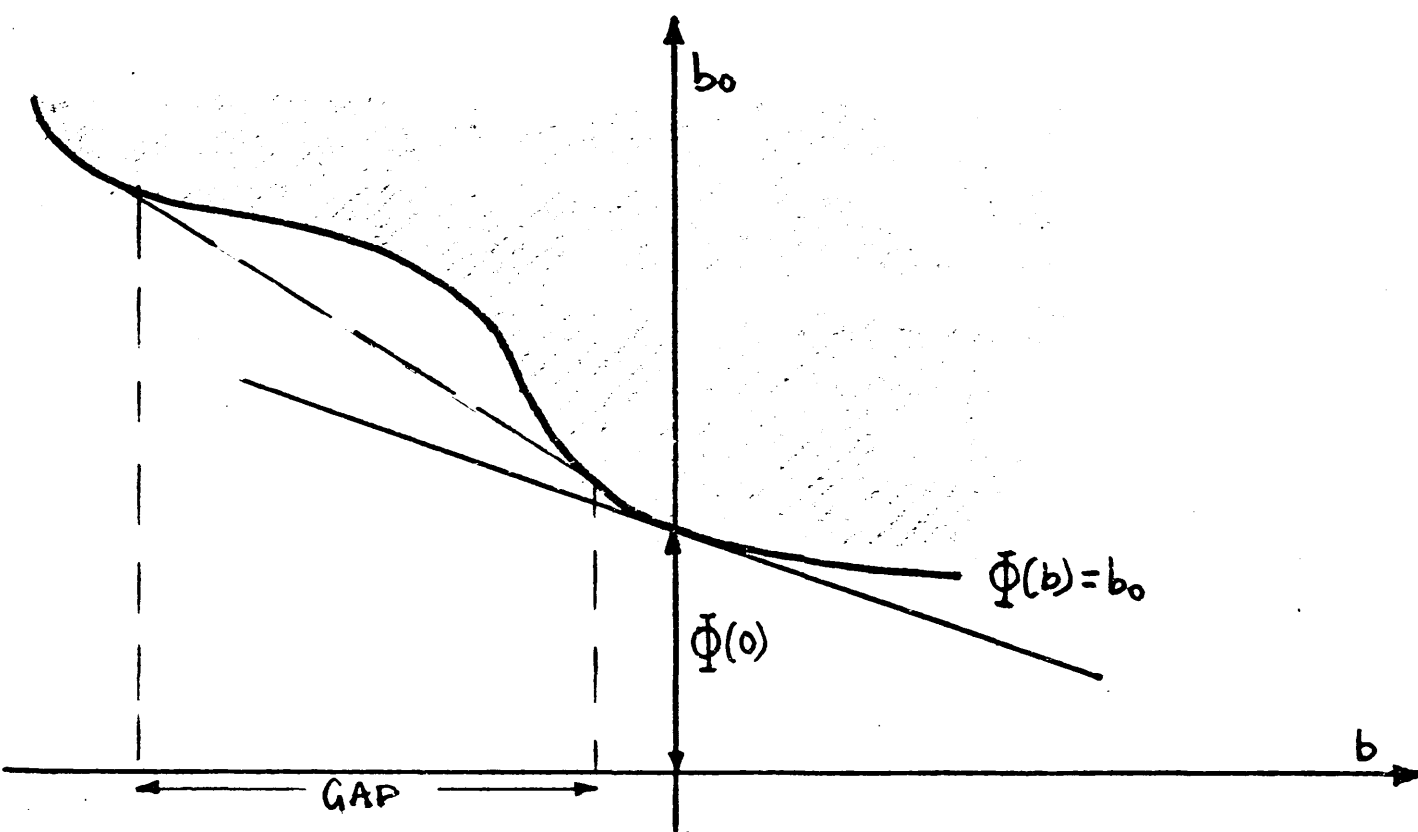
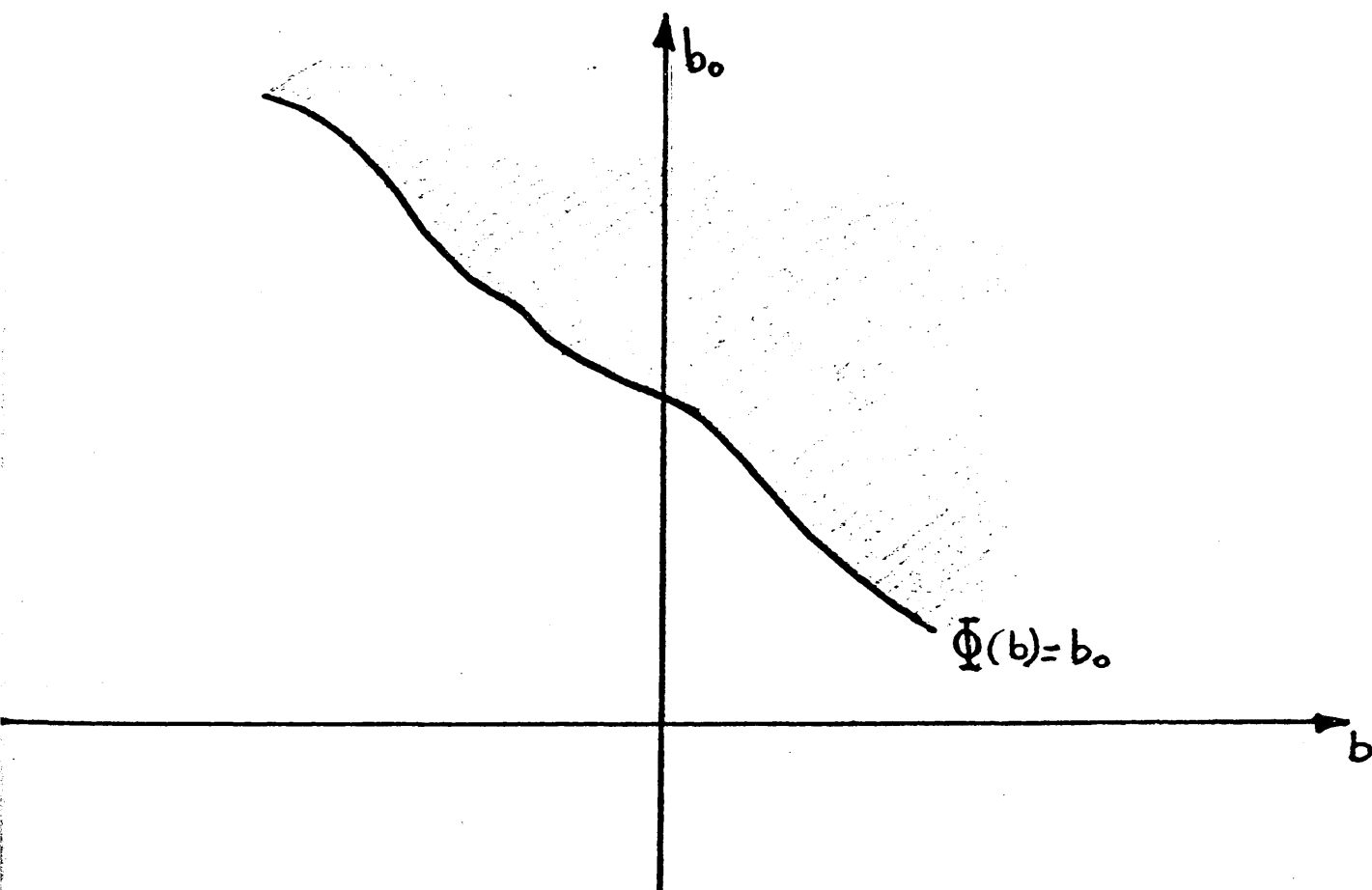


FIGURA EV.16 PROBLEMA NO CONVEXO NO RESOLVIBLE POR METODOS
DUALES DE DESCOMPOSICION JERARQUICA



Si las condiciones de convexidad son objeto de especial estudio es debido fundamentalmente a tres series de razones:

- a) Cuando se supone o acepta la convexidad del problema, se pueden derivar importantes resultados teóricos.
- b) Los resultados obtenidos cuando se suponen reunidas las condiciones de convexidad permiten recoger información acerca de las propiedades de problemas no convexos.
- c) Las propiedades de programas convexos pueden extenderse en formas más débiles, a problemas no convexos.

El mismo tipo de razones puede justificar el desarrollo de métodos y modelos basados en la hipótesis de la existencia de hiperplanos soportes del conjunto C en el punto solución. Puesto que hemos visto el interés de los resultados obtenidos, aún en el caso de que esta hipótesis se revelara falsa a posteriori, creemos que los métodos de descomposición jerárquica en la solución multinivel de problemas complejos de control óptimo son, a la vez, del mayor interés - y de las pocas alternativas susceptibles de generar métodos operativos en problemas complejos de elevadas dimensiones.

Como ha quedado expuesto, las posibilidades de obtener la solución descentralizada del problema, dependerán de las características de este.

Como conclusión a la discusión de este tema, es de interés contrastar la estructura obtenida en el modelo de resolución descentralizada del problema de control, con los comentarios referentes a las posibilidades de conseguir este tipo de soluciones expuestas por Armand (A8) (referidas en

CAPITULO VI

CONCLUSIONES

Say not "I have found the truth", but "I have found a truth"

Kahlil Gibsan

A lo largo de los anteriores capítulos, se han ido explicitando los presupuestos que los motivan:

- a) Las decisiones económicas dependen de preferencias políticas que, una vez definidas, deben ser consideradas como datos en modelos matemáticos dinámicos de la realidad a la que se refieren. De la solución de los mismos pueden obtenerse valiosas indicaciones acerca de la forma óptima de satisfacer tales preferencias, que escapan a otros tipos de análisis.
- b) Los valores de los parámetros de los modelos económicos y los datos que los alimentan se conocen con un elevado grado de imprecisión. Por ello:
 - b1) más que en un problema específico, el interés radica en una familia de problemas de la misma estructura que cubran una gama de alternativos valores de sus parámetros.
 - b2) la información dual es tan importante como la solución del propio problema primal.
- c) La formulación de complejos modelos matemáticos debe ir acompañada de un esfuerzo considerable en los algoritmos y métodos numéricos de resolución, requisito indispensable para conseguir una operatividad que los justifique. Por ello:
 - c1) es preciso desarrollar procedimientos eficientes que permitan abordar la solución de problemas reales, que se caracterizan por:
 - su carácter dinámico y elevada dimensionalidad

- poseer estructuras específicas que pueden facilitar el proceso de cálculo

c2) los métodos empleados deben facilitar la información dual de forma fiable y obtenerla sin un coste excesivo de cálculo suplementario.

c3) los métodos empleados deben ser capaces de resolver modelos no lineales y con restricciones instantáneas sobre controles y estados.

d) Los modelos matemáticos deben reflejar la existencia real de varios centros de decisión.

En el contexto así definido, consideramos que la moderna teoría matemática del control óptimo y sus desarrollos subsiguientes en términos de juegos diferenciales y sistemas jerárquicos de control multinivel, constituyen, tanto numérica como conceptualmente, un marco instrumental de decisiva importancia en la formulación, resolución y análisis de modelos de planificación económica.

En el capítulo II hemos presentado, de forma unificada, el aparato interpretativo en términos económicos asociado a los elementos matemáticos de dicha teoría. Se han analizado (II. 2-7) los distintos significados asociados a las variables adjuntas en el marco general del análisis de sensibilidad de un sistema dinámico frente a las distintas perturbaciones que pueden afectarle. La forma en que tal significado se extiende al caso en el que existen varios centros de decisión, se ha presentado (II. 8) mediante el análisis de las trayectorias de Nash en la teoría de juegos diferenciales.

Los problemas de cálculo asociados a los modelos econó-

micos dinámicos del tipo que formula la Teoría del control, se han considerado en el capítulo III. El análisis de las alternativas entre formulaciones discretas y continuas, y entre los métodos directos e indirectos de resolución (III.2) nos ha conducido al estudio comparativo de los tipos primales y duales de los algoritmos de programación matemática con los que se puede tratar la resolución de problemas discretos no lineales de control óptimo.

Las ventajas e inconvenientes respectivos de ambos tipos han sido analizados. Primero de una forma global (III.2), y después mediante el estudio detallado de un método primal (III.3) y dos métodos duales (III.4).

Los métodos duales requieren determinadas condiciones de convexidad en el problema para ser aplicables. Su proceso de cálculo es relativamente sencillo y especialmente ilustrativo económicamente. No mantienen la factibilidad de las soluciones del proceso iterativo pero no exigen el conocimiento previo de una solución factible inicial.

Los métodos primales son universalmente aplicables, aunque, por supuesto, no garanticen sino un óptimo local en problemas no convexos. Mantienen la factibilidad de todas las soluciones a costa de un proceso de cálculo numéricamente muy complejo y el conocimiento previo de una solución factible inicial.

Este último requisito es especialmente difícil de satisfacer cuando el problema de control es muy rico en restricciones instantáneas mixtas control/estado. Ello constituye una primera y considerable ventaja a favor de los métodos duales.

Por otra parte, la formulación de un problema de control óptimo discreto como un programa matemático (III.2) resulta en uno de tan elevadas dimensiones que la única posibilidad de resolverlo estriba en métodos que aprovechen su estructura dinámica, especialmente representada a través del Jacobiano de las restricciones (III.3.3, III.4.2).

Los métodos primales no producen ningún tipo de descomposición del problema. Una de las mayores ventajas de la aplicación de un algoritmo GRG (III.3.1,2) a un problema de control (III.3.3) es la posibilidad de descomponer secuencialmente el cálculo del gradiente reducido, lo cual equivale a calcular, secuencialmente y sin coste suplementario, las variables adjuntas. La dimensión de las matrices a invertir queda así reducida a valores operativos. En III.3.3 damos un organigrama detallado de este proceso de cálculo, donde queda patente, como analizamos en III.3.4, que no es automáticamente aplicable a todos los problemas.

Los métodos duales, por el contrario, descomponen el problema en dos niveles de cálculo iterativo (III.4.1). En algunos problemas, puede además descomponerse el problema del primer nivel (problema Lagrangiano) en una serie de problemas locales, de los que el segundo nivel (función dual) actúa como centro coordinador.

Si la aplicabilidad de los métodos duales depende de que el problema satisfaga la hipótesis de convexidad local (III.4.1), la descomponibilidad del problema del primer nivel depende de la estructura aditiva y separable de la función objetivo y restricciones del problema primal.

Un problema de control posee tal estructura. Por ello,

la aceptación de la hipótesis de convexidad local permite la aplicación de algoritmos tipo Lasdon-Tamura (III.4)..Su utilización permite tratar problemas con efectos retardados en controles y estados, sin aumentar las dimensiones del vector de estado (III.4).

En problemas que no posean tal propiedad, el empleo del método de Héstenes para modificar el enunciado del problema permite dotarle de la misma. En III.4.2 construimos un método dual mediante la introducción de tal modificación. Damos el organigrama del algoritmo resultante y analizamos su aplicabilidad real (requerimientos en memoria rápida y cálculo) en función de los retrasos en el tiempo que presente la formulación del problema y del número de restricciones instantáneas que incluya. Mostramos la viabilidad del mismo, así como la especial sencillez de su proceso de cálculo, que no exige inversiones ni soluciones de sistemas de ecuaciones no lineales. Su mayor inconveniente es el de destruir la descomposición del problema del primer nivel.

Todos los algoritmos analizados son de carácter global, es decir, simulan un único centro estructural de decisión en el sistema. Solamente el método dual puro de III.3.4 descompone temporalmente el problema en una serie de centros de decisión instantáneos, de acuerdo con el principio del Máximo (Apéndices I y II), y con los Hamiltonianos por este definidos actuando de objetivos locales de cada uno de ellos.

Una vez analizados los condicionamientos que el proceso de cálculo impone, hemos estudiado (capítulo IV) la formulación de modelos dinámicos multisectoriales de planificación, como problemas de control óptimo discreto. Del análisis de las ventajas e inconvenientes del uso del sistema lineal-cua-

drático en la construcción de este tipo de modelos, concluimos:

- la aplicación de las "reglas de decisión lineales" es mucho menos eficiente que la de los métodos variacionales

- el uso del modelo lineal-cuadrático es especialmente indicado para fines de regulación, que son de naturaleza distinta a los propósitos de un proceso planificador a largo plazo.

- en un proceso a largo plazo deben evitarse especialmente las hipótesis de linealidad

- el modelo dinámico de input-output debe descomponerse en tres grupos de relaciones, que representan respectivamente los procesos de producción, distribución y capitalización. De ellos, solamente el segundo se considera lineal

- el cálculo de la trayectoria nominal precisada para la aplicación del modelo lineal-cuadrático debe efectuarse mediante modelos no lineales.

En el análisis del modelo no lineal de Kendrick y Taylor (K6 , A1), hemos considerado las siguientes limitaciones en su formulación:

- hipótesis de pleno empleo
- posibilidad ilimitada de empleo productivo del factor trabajo con independencia de las dotaciones de capital
- niveles instantáneos no acotados de deuda exterior

Tan irreales hipótesis son impuestas por la no consideración de restricciones instantáneas sobre la trayectoria. Las restricciones del tipo $a \leq x(t) \leq b$ (Abadie (A1)) no solucionan sino la última de tales limitaciones. La superación de las dos primeras, exige la formulación de restricciones instantáneas mixtas control/estado. Ello limita fuertemente la aplicabilidad de los métodos primales, por las difi-

cultades asociadas al previo conocimiento de una solución inicial factible.

Presentamos un planteamiento alternativo que incluye:

- el abandono de la hipótesis de pleno empleo y la valoración negativa del mismo
- la limitación de las posibilidades de empleo con cantidades dadas de capital, mediante una restricción del tipo $g(l_i(t), k_i(t), \partial q_i(t)/\partial l_i(t)) \leq 0$
- la acotación de las posibilidades instantáneas de deuda exterior.

Hemos considerado el modelo así constituido (IV.3.3), como un problema de control con centro de decisión único, y analizado las posibilidades de su resolución, de lo que concluimos:

- la solución por el método primal GRG (III.3) y el "dual de los multiplicadores" (III.4.2) exige la consideración del mismo número de variables y restricciones
- ninguno de ellos presupone la convexidad de la función objetivo
- el método dual de los multiplicadores es especialmente favorable para su aplicación numérica
- ninguno de ellos es susceptible de aplicación eficiente con niveles de desagregación elevados.

Los únicos procedimientos que permiten tratar modelos reales son los de descomposición-descentralización, impuestos por:

- la necesidad de considerar centros de decisión locales, que están realmente presentes en la estructura que el modelo representa
- los límites en la viabilidad práctica de los métodos de solución global (IV.4).

Consideramos que la vía alternativa de modelos de planificación basada en la Teoría de los juegos diferenciales y desarrollada por Pau (P3), es más apropiada para modelos de economías de mercado. En economías socialistas planificadas, los métodos de descomposición jerárquica multinivel de problemas de control son la técnica más adecuada para analizar el funcionamiento independiente de las partes dentro del conjunto y permitir una solución numérica del problema (IV.1,2).

En V.3 hemos planteado la solución descentralizada del modelo de planificación de IV.3, demostrando la estructura descomponible del problema del primer nivel y explicitando las relaciones de interconexión entre sus centros locales.

El proceso resultante es absolutamente descentralizado, en el sentido que:

- la autoridad del primer nivel no genera propuestas primales
- los planes de actuación son fijados total y exclusivamente por los centros locales del primer nivel

Sus características son las siguientes:

- la autoridad coordinadora del segundo nivel tiene únicamente a su cargo la gestión de las restricciones de interconexión
- dicha autoridad debe fijar únicamente los precios de los bienes sobre los que pesan tales restricciones
- los segundos miembros de las restricciones de interconexión referentes a la disponibilidad del factor trabajo y del nivel máximo de deuda exterior son los únicos que, por su carácter de magnitudes exógenas, son conocidos a priori por dicha autoridad

- los restantes segundos miembros son fijados dinámicamente por el conjunto de decisiones locales de los centros, y no son conocidos por la autoridad coordinadora al principio de cada iteración
- la autoridad coordinadora precisa conocer una información muy reducida acerca de los conjuntos de producción de los centros. En particular, no precisa conocer sus funciones de producción en lo que a inputs primarios se refiere
- cada centro local del primer nivel (autoridad responsable de cada sector) debe resolver, en cada iteración, un problema autónomo de control óptimo, bajo el sistema de precios de sus inputs y outputs generado por la autoridad coordinadora
- los centros deben generar una información exterior muy limitada, toda ella del tipo primal.
- los precios de los bienes de capital son generados por y utilizados en el subsistema correspondiente exclusivamente, no precisando ser conocidos por el centro coordinador
- la formulación sitúa todas las no linealidades en los problemas locales del primer nivel, mientras que genera una forma lineal de las restricciones de interconexión. Ello es especialmente interesante, tanto para la aplicabilidad del método como por la estructura de los datos disponibles para alimentarlo.

El método genera un óptimo global en problemas convexos, al mismo tiempo que produce una información valiosísima para un análisis de sensibilidad de las soluciones.

La convexidad no es una condición necesaria para la aplicabilidad del método. Es posible caracterizar problemas no convexos que pueden ser resueltos eficazmente mediante el uso de los mismos. La existencia de un hiperplano soporte en puntos de un determinado conjunto C es la condición precisada (Lasdon((L3))). Ello no constituye un test de aplicabilidad a priori, pero tampoco existe un procedimiento fácilmente aplicable para determinar la convexidad de un problema.

Utilizando los teoremas de Everett (E2), hemos demostrado como, de acuerdo con los presupuestos enunciados, la información generada a lo largo del proceso de cálculo es de tanta o mayor importancia como la solución del problema planteado: Ello es independiente de que se llegue o no a alcanzar la solución del mismo.

En resumen:

- La introducción de hipótesis realistas impone la formulación de problemas de control no lineales con restricciones instantáneas mixtas.
- Ello dificulta la aplicación de los métodos de tipo primal.
- Los métodos primales no generan ninguna descomposición del problema. El método de Hestenes induce un algoritmo dual a dos niveles que garantiza la convexidad del problema modificado pero destruye la descomposición del problema del primer nivel.
- La única posibilidad de resolver problemas de dimensiones reales es instrumentar procedimientos de descomposición-descentralización del problema global de control.

- La aplicación de los métodos de descomposición jerárquica multinivel tipo Lasdon-Tamura exige la satisfacción por el problema de determinadas condiciones, menos restrictivas que las de su convexidad.
- Si tales condiciones se verifican, se obtiene un procedimiento de resolución a dos niveles completamente descentralizado.
- Aunque tales condiciones no se verifiquen, y por lo tanto no se obtenga la solución del problema propuesto, la información generada en el proceso de cálculo justifica por si sola, la aplicación del método.

El proceso de cálculo tiene, a lo largo de todas las iteraciones, una rica interpretación económica. Ello nos permite esperar no merecer ser incluidos en la cita con la que encabezábamos este trabajo.

APENDICES

NOMENCLATURA DE UN PROBLEMA CONTINUO DE CONTROL OPTIMO Y
FORMULACION GENERAL DEL PRINCIPIO DEL MAXIMO DE PONTRYAGIN.

Dado un sistema dinámico que evoluciona entre un instante inicial 0 y un instante final T , cuyo estado esta representado en cada instante por las n componentes de su VECTOR DE VARIABLES DE ESTADO $x(t)$:

$$X(t) = \begin{bmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{bmatrix}$$

supuestas continuas en $[0, T]$

Sobre la evolución del sistema es posible actuar mediante las m VARIABLES DE CONTROL O DE DECISION $\mu(t)$.

$$\mu(t) = \begin{bmatrix} \mu_1(t) \\ \vdots \\ \mu_m(t) \end{bmatrix}$$

supuestas continuas por intervalos en $[0, T]$, es decir, continuas $\forall t \in [0, T]$ excepto en un número finito de valores de t en los que pueden tener discontinuidades de primer orden.

Los valores que pueden tomar las variables de control son los pertenecientes al conjunto definido por las RESTRICCIONES SOBRE LOS CONTROLES :

$$\mu(t) \in U(t) \in E_m$$

La dinámica del sistema, supuesta determinista, está de-

finida por el sistema de n ecuaciones diferenciales o

ECUACIONES DE ESTADO

$$\dot{x}(t) = f(x(t), u(t), t)$$

donde $f(x(t), u(t), t)$ es un n -vector de funciones

$$f = \begin{bmatrix} f_1 \\ \vdots \\ f_n \end{bmatrix}$$

continuas con respecto a x, u y continuamente diferenciables con respecto a x .

A cada posible trayectoria del sistema se le asigna un valor determinado por la FUNCIONAL OBJETIVO :

$$J = \Phi(x(T)) + \int_0^T f_0(x(t), u(t), t) dt$$

donde la función f_0 tiene las mismas características enunciadas para las funciones f .

El problema consiste en determinar la trayectoria factible de las variables de control $u^*(t), \forall t \in [0, T]$ (llamada CONTROL OPTIMO DEL SISTEMA), de forma que la trayectoria seguida por el sistema como consecuencia de su aplicación, a partir de un cierto ESTADO INICIAL $x(0)$, sea tal que la funcional objetivo alcance su máximo.

El problema se resume pues en :

$$\begin{aligned} & \underset{\substack{u(t) \\ \forall t \in [0, T]}}{\text{Max}} \left[\Phi(x(T)) + \int_0^T f_0(x(t), u(t), t) dt \right] \\ & \dot{x}(t) = f(x(t), u(t), t) \end{aligned}$$

$$X(0) = X_0$$

$$u(\tau) \in U(\tau) \in E_m$$

El valor de T , INSTANTE FINAL del proceso, puede ser fijo y constituir un dato del problema, o puede ser indeterminado y su valor óptimo formar parte del resultado.

Este Apéndice se propone simplemente resumir estructuralmente la terminología y principales resultados de la Teoría del Control Óptimo. El lector no familiarizado con esta teoría debe dirigirse a un más extenso tratamiento de la misma en la obra original de Pontryagin (P12) o en la numerosa bibliografía existente (A7, H12, H2)

CONDICIONES EN LOS LIMITES

Sobre el estado inicial $x(0)$ y final del sistema $x(T)$ pueden ser impuestos distintos condicionamientos:

CONDICIONES INICIALES

Suponemos que el estado inicial es conocido y fijo

$$X(0) = X_0$$

Matemáticamente nada obliga a formular dicha hipótesis, pero es la generalmente impuesta por la realidad en las aplicaciones económicas.

CONDICIONES FINALES

Distintos casos deben ser considerados.

- a) Todas, o algunas variables de estado libres en T :

b) Restricciones sobre el estado final

$$X(T) \in S \subset E_n$$

El dominio S está definido, en su forma más general, por las restricciones desigualdad :

$$M(X(T), T) \geq 0$$

donde M es un vector de funciones 1-dimensional

$$M = \begin{bmatrix} M_1 \\ \vdots \\ M_e \end{bmatrix}$$

Caso particular interesante es aquel en el que todas, o parte, de las variables de estado en T están fijadas a un cierto valor:

$$M_i(X(T), T) = X_i(T) - X_{i,T} = 0 \quad (i=1, \dots, p \leq n)$$

Con esta notación, las condiciones enunciadas por el PRINCIPIO DEL MAXIMO DE PONTRYAGIN son :

- Para que un vector admisible de variables de decisión $\mu^*(t)$ y la correspondiente trayectoria $X^*(t)$ sean óptimos, es necesario que exista un vector no nulo de VARIABLES ADJUNTAS $\Psi(t)$, continuas y diferenciables por intervalos :

$$\Psi(t) = \begin{bmatrix} \Psi_1(t) \\ \vdots \\ \Psi_n(t) \end{bmatrix}$$

que obedezcan al sistema de ecuaciones diferenciales:

$$\boxed{\dot{\Psi}_i(t) = -\frac{\partial H}{\partial x_i}} \quad (i=1, \dots, n)$$

con las siguientes condiciones en los límites :

1) CONDICION DE TRANSVERSALIDAD :

$$\Psi'(T) = \nabla_{x(T)} [\Phi(x^*(T), T)] + \lambda' \nabla_{x(T)} [M(x^*(T), T)]$$

2) CONDICIONES DE IMPACTO EN EL BLANCO :

$$\begin{aligned} \lambda' M(x^*(T), T) &= 0 \\ \lambda &\geq 0 \end{aligned}$$

en el que la función HAMILTONIANO se define :

$$H[x(t), u(t), \psi(t), t] = f_0(x, u, t) + \sum_{i=1}^n \psi_i(t) f_i(x, u, t)$$

y tales que en todo $t \in [0, T]$ el HAMILTONIANO, considerado como una función de $u(t)$, alcance su máximo para $u(t) = u^*(t)$

$$\begin{aligned} \max_{u \in U} [H[x(t), u(t), \psi(t), t]] &= \\ = H[x^*(t), u^*(t), \psi(t), t] &= H^*[x(t), u(t), \psi(t), t] \end{aligned}$$

maximización en la que $\psi(t), x(t)$ se consideran como parámetros.

Si el valor de T no esta fijado, una CONDICION DE TRANSVERSALIDAD suplementaria determina su valor óptimo:

$$\left[\frac{\partial \Phi}{\partial t} + \lambda' \nabla_t M + H^* \right]_{t=T} = 0$$

PARTICULARIZACIONES

Distintas circunstancias especiales en la definición del sistema conducen a resultados de especial interés :

CONDICIONES EN LOS LIMITES

- 1- Si no existe función de valoración asociada con el estado final :

$$\Phi(x(\tau), \tau) \equiv 0$$

y si el dominio del espacio E_n al que debe pertenecer $x(\tau)$ está definido como la intersección de l hipersuperficies :

$$M_i(x(\tau)) = 0 \quad (i=1, \dots, l)$$

La CONDICION DE TRANSVERSALIDAD adopta la forma originalmente enunciada por Pontryagin :

$$\psi'(\tau) = \lambda' \nabla_{x(\tau)} [M(x^*(\tau))]$$

Expresión que indica que el vector $\psi(\tau)$ es ortogonal al hiperplano tangente en $x(\tau)$ a la hipersuperficie definida por :

$$M(x(\tau)) = 0$$

- 2- Si el estado final está completamente fijado :

$$M_i(x(\tau)) \equiv x_i(\tau) - x_{i,\tau} = 0 \quad (i=1, \dots, n)$$

el vector $\psi(\tau)$ queda indeterminado :

$$\psi(\tau) = \lambda$$

- 3- En ausencia de todo tipo de restricciones y de valor de fin de proceso :

$$\boxed{\psi(\tau) = 0}$$

expresión que constituye la llamada condición natural" en los límites.

- 4- Si no existe determinación alguna del estado final pero si valor de fin de proceso: $\Phi(x(T), T)$:

$$\psi(T) = \nabla_{x(T)} [\Phi(x^*(T), T)]$$

- 5- Todas estas consideraciones pueden extenderse al instante inicial:

$$\psi_i(0) = 0 \quad \text{si} \quad x_i(0) \quad \text{es completamente libre}$$

$$\psi_i(0) \quad \text{es indeterminado si} \quad x_i(0) = x_{i,0}$$

DETERMINACION DE T

- 1- La condición:

$$\left[\frac{\partial \Phi}{\partial t} + \lambda' \nabla_t M + H^* \right]_{t=T} = 0$$

se reduce a:

$$H^*_{(t=T)} = 0$$

si las restricciones sobre el estado final y la función de valor de fin de juego son independientes de t o inexistentes.

- 2- La propiedad del Hamiltoniano óptimo :

$$\frac{\partial H^*}{\partial t} = \frac{dH^*}{dt}$$

implica su CONSTANCIA a lo largo del proceso para un proceso autónomo (estacionario):

$$\forall t \in [0, T], \quad H^*(t) = \dot{c} = H^*(T) = 0$$

Como consecuencia, la ecuación :

$$H^*(t) = 0$$

aparece como una integral primera del sistema definiendo $\dot{\Psi}$ y \dot{X} , y ya no es susceptible de permitir el cálculo de T por si misma.

3- En el caso de un sistema no autónomo :

$$\frac{dH^*}{dt} = \frac{\partial H^*}{\partial t} = \sum_{i=0}^n \frac{\partial f_i(x(t), u(t), t)}{\partial t} \psi_i(t)$$

habiendo definido:

$$\psi_0(t) \equiv 1$$

y puesto que, conservando las mismas hipótesis sobre el estado final :

$$H^*(T) = 0$$

aparece:

$$H^*(t) = - \int_t^T \sum_{i=0}^n \frac{\partial f_i(x, u, t)}{\partial t} \psi_i(t) dt$$

4- En los problemas de tiempo óptimo, en los que $f_0 \equiv 1$, $\Phi \equiv 0$, necesariamente debe de existir un conjunto de restricciones sobre el estado final. La condición de transversalidad en T , teniendo en cuenta que:

$$\Psi'(T) = \lambda' \nabla_{x(T)} [M(x^*(T), T)]$$

se expresa:

$$\boxed{\left[\lambda' \nabla_t M + [\lambda' \nabla_x M] f + 1 \right]_{t=T} = 0}$$

Según la definición de las VARIABLES ADJUNTAS $\psi(t)$ y del HAMILTONIANO :

$$\begin{aligned}\dot{\psi}_i(t) &= -\frac{\partial H}{\partial x_i} \\ \dot{x}_i(t) &= \frac{\partial H}{\partial \psi_i}\end{aligned}$$

$$i = (1, \dots, n)$$

este sistema de $2n$ ecuaciones diferenciales es llamado SISTEMA CANONICO.

Precisa de $2n$ condiciones en los límites, que vienen dadas por:

Las primeras n están dadas por las condiciones en $t=0$:

$$\psi_i(0) = 0 \quad \text{o bien} \quad x_i(0) = x_{i,0}$$

Las n restantes se definen en función de las condiciones definidas para el vector estado en T . En general, a través de las condiciones de transversalidad, aparecen l constantes indeterminadas λ , a las que, en algunos casos, se les añade el valor desconocido de T .

La determinación de las $n+l+1$ constantes, exige que las condiciones en T faciliten $n+l+1$ ecuaciones, de las cuales n son definidas por la condición de transversalidad (1), l por las condiciones de impacto en el blanco (2) y 1 por la condición de transversalidad en T (3).

En el caso elemental en el que el problema define como datos los valores de T y de las variables de estado en el instante final $x(T)$, el sistema canónico se reduce a un problema de contorno de orden $2n$ con las condiciones

en los límites distribuidas entre 0 y T:

$$X_i(0) = x_{i,0} \text{ o bien } \psi_i(0) = 0$$

$$X_i(T) = x_{i,T} \text{ o bien } \psi_i(T) = 0$$

que, al menos teóricamente, bastan para determinar las $2n$ constantes de integración.

NOMENCLATURA DE UN PROBLEMA DISCRETO DE CONTROL OPTIMO
Y FORMULACION DEL PRINCIPIO DEL MAXIMO DISCRETO.

Si el problema descrito en el Apendice I se formula en TIEMPO DISCRETO, las ecuaciones diferenciales que describen la evolución del sistema se convierten en ecuaciones en diferencias finitas, y el objetivo se expresa como un sumatorio extendido a todos los instantes de tiempo considerados.

Correlativamente a como se planteaba en el Apéndice I, y utilizando la misma nomenclatura, el problema formulado en tiempo discreto es el de:

$$\underset{\substack{\mu(k) \\ k=0, \dots, N-1}}{\text{Max}} \left[\Phi(x(N)) + \sum_{k=0}^{N-1} f_0(x(k), \mu(k), k) \right]$$

$$x(k+1) = f(x(k), \mu(k), k) \quad (k=0, \dots, N-1)$$

$$g(x(k), \mu(k), k) \geq 0 \quad (k=0, \dots, N-1)$$

$$x(0) = x_0$$

Los elementos que intervienen en su enunciado son:

FUNCION OBJETIVO

$$J = \Phi(x(N)) + \sum_{k=0}^{N-1} f_0(x(k), \mu(k), k)$$

$$X(k+1) = X(k) + f(X(k), u(k), k) \quad k=0, \dots, T-1$$

RESTRICCIONES

Adoptan la forma general:

$$g(X(k), u(k), k) \geq 0 \quad k=0, \dots, T-1$$

CONDICIONES EN LOS LIMITES

ESTADO INICIAL supuesto fijo :

$$X(0) = X_0$$

ESTADO FINAL perteneciente a un cierto dominio definido por :

$$g^N(X(N)) \geq 0$$

que admite como caso particular el de estado final fijo:

$$g^N(X(N)) \equiv X(N) - X_N = 0$$

Donde:

$X(k)$ es el VECTOR DE ESTADO n-dimensional.

$$X(k) = \begin{bmatrix} x_1(k) \\ \vdots \\ x_n(k) \end{bmatrix}$$

$u(k)$ es el VECTOR DE DECISION O DE CONTROL m-dimensional.

$$u(k) = \begin{bmatrix} u_1(k) \\ \vdots \\ u_m(k) \end{bmatrix}$$

$g(x(N))$ son cóncavas, las funciones $f_i(x, u, k)$ lineales en u , y todas ellas diferenciables.

Considerado como un programa matemático, el problema anterior adopta la forma :

$$\text{Max}_{[u(0), \dots, u(k), \dots, u(N-1), x(0), \dots, x(N)]} \left[\Phi(x(N)) + \sum_{k=0}^{N-1} f_0(x(k), u(k), k) \right]$$

bajo las restricciones :

a) de tipo igualdad

$$x(0) = x_0$$

$$x(k+1) - x(k) = f(x(k), u(k), k) \quad (k=0, N-1)$$

b) de tipo desigualdad

$$g(x(k), u(k), k) \geq 0 \quad (k=0, N-1)$$

$$g^N(x(N)) \geq 0$$

Llamando:

$$\psi(0) = \begin{bmatrix} \psi_1(0) \\ \vdots \\ \psi_n(0) \end{bmatrix}$$

el vector de variables duales asociado a $x(0) = x_0$.

$$\psi(k+1) = \begin{bmatrix} \psi_1(k+1) \\ \vdots \\ \psi_n(k+1) \end{bmatrix}$$

..

..

$$\begin{aligned} \text{.. } x(k+1) - x(k) &= \\ &= f(x(k), u(k), k) \\ &(k=0, N-1) \end{aligned}$$

$$\lambda(k) = \begin{bmatrix} \lambda_1(k) \\ \vdots \\ \lambda_e(k) \end{bmatrix}$$

≥ 0 ..

..

$$\begin{aligned} \text{.. } g(x(k), u(k), k) &\geq 0 \\ &(k=0, N-1) \end{aligned}$$

$$\lambda(\tau) = \begin{bmatrix} \lambda_1(n) \\ \vdots \\ \lambda_e(n) \end{bmatrix} \quad " \quad " \quad " \quad " \quad " \quad " \quad g^N(x(n)) \geq 0$$

y utilizando la simplificación notacional:

$$\left. \begin{aligned} g^N(x(N)) &= g^*(N) \\ g(x(k), u(k), k) &= g^*(k) \\ f(x(k), u(k), k) &= f^*(k) \\ f_0(x(k), u(k), k) &= f_0^*(k) \end{aligned} \right\}$$

Las condiciones de Kuhn y Tucker referentes al Lagrangiano : **N-1** **N-1**

$$\mathcal{L} = \Phi(x(N)) + \sum_{k=0}^{N-1} f_0(k) + \sum_{k=0}^{N-1} \psi'(k+1) [f(k) + x(k) - x(k+1)] + \psi'(0) [x(0) - x_0] + \sum_{k=0}^{N-1} \lambda'(k) g(k) + \lambda'(N) g(N)$$

se escriben :

Con respecto a las variables $x(k)$, para los distintos valores de k , ($k = 0, \dots, N$):

$$k=0, N-1 \left\{ \begin{array}{l} \nabla_{x(k)} [f_0^*(k)] + \psi'(k+1) \nabla_{x(k)} [f^*(k)] + \psi'(k+1) - \psi'(k) + \\ + \lambda'(k) \nabla_{x(k)} [g^*(k)] = 0 \\ x^*(k+1) - x^*(k) = f(x^*(k), u^*(k), k) \\ \lambda'(k) g^*(k) = 0 \end{array} \right.$$

$$K=N \quad \begin{cases} \nabla_{x(N)} [\Phi] + \lambda'(N) \nabla_{x(N)} [g^*(N)] - \psi(N) = 0 \\ \lambda'(N) g^*(N) = 0 \end{cases}$$

$$K=0 \quad x(0) = x_0$$

Con respecto a las variables $u(k)$, para los distintos valores de $k, (k=0, N-1)$

$$K=0, \dots, N-1 \quad \begin{cases} \nabla_{u(k)} [f_0^*(k)] + \psi'(k+1) \nabla_{u(k)} [f^*(k)] + \\ + \lambda'(k) \nabla_{u(k)} [g^*(k)] = 0 \\ \lambda'(k) g^*(k) = 0 \end{cases}$$

Definiendo la función HAMILTONIANO en el instante K :

$$H(k) = f_0(x(k), u(k), k) + \psi'(k+1) f(x(k), u(k), k)$$

y substituyendo su expresión en las anteriores condiciones de Kuhn y Tucker, estas pueden expresarse y agruparse de la siguiente forma:

$$\begin{aligned} \psi'(N) &= \nabla_{x(N)} [\Phi] + \lambda'(N) \nabla_{x(N)} [g^*(N)] \\ \psi'(k+1) - \psi'(k) &= -\nabla_{x(k)} [H(k)] - \lambda'(k) \nabla_{x(k)} [g^*(k)] \\ x^*(k+1) - x^*(k) &= f(x^*(k), u^*(k), k) \end{aligned} \quad \boxed{I}$$

$$\lambda'(k) g^*(k) = 0$$

$$\nabla_{\mu(k)} [H(k)] + \lambda'(k) \nabla_{\mu(k)} [g^*(k)] = 0$$

$$(k=0, N-1)$$

II

$$X(0) = X_0$$

$$\lambda'(N) g^*(N) = 0$$

III

que constituyen el enunciado del PRINCIPIO DEL MAXIMO DISCRETO.

Las Ecuaciones II representan las Condiciones de Maximización del Hamiltoniano en cada instante k con respecto a las solas variables de decisión de este instante $\mu(k)$. Expresan pues la búsqueda de una serie de óptimos estáticos de la nueva función de evaluación que el Hamiltoniano representa.

Las ecuaciones I elaboran la información necesaria para la construcción del Hamiltoniano. Las expresadas en III enuncian las condiciones que la trayectoria debe satisfacer en los instantes inicial y final.

El resultado puede ser anunciado de la siguiente forma:

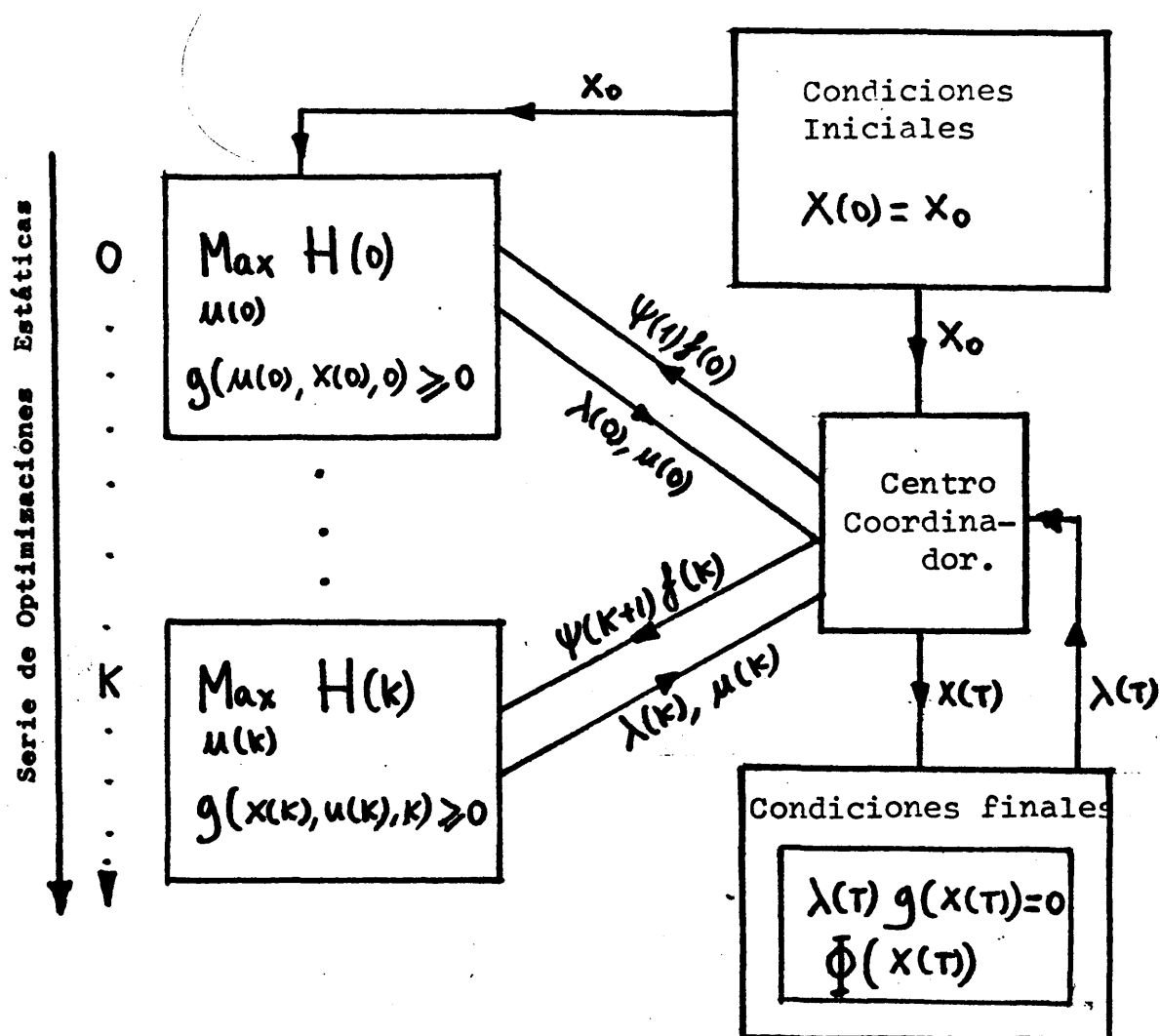
- El óptimo dinámico puede ser descompuesto en una serie equivalente de óptimos instantáneos de la nueva función de evaluación:

$$H(k) = f_0(k) + \psi'(k+1) f(k)$$

en la que $\Psi(k)$ y $\lambda(k)$ obedecen las ecuaciones en diferencias finitas descritas en I .

Observese que el proceso descentralizado de optimización precisa de un centro coordinador que resuelva las ecuaciones I . Ello implica un flujo de información como el esquematizado por Albouy (17) y representado en la figura.

Este intercambio de información primal-dual entre un centro coordinador y los centros autónomos encargados de los procesos locales de optimización es de la mayor importancia en las aplicaciones económicas y será ampliamente referido en los capítulos II, III y V.



APENDICE III.

METODOS DE OPTIMIZACION NO LINEAL.

A. SIN RESTRICCIONES.

Los métodos iterativos básicos son de la forma general:

$$x_{k+1} = x_k \pm \bar{\alpha}_k S_k \nabla' f(x_k)$$

$$\bar{\alpha}_k = \arg \max_{\alpha} \min_{\alpha} f(x_k \pm \alpha S_k \nabla' f(x_k))$$

El método del gradiente se obtiene con $S_k = I$. Converge globalmente pero lo hace linealmente y con razón de convergencia determinada por el cociente de los autovalores máximo y mínimo del Hessiano del objetivo en el óptimo (razón canónica). Dependiendo de la estructura de los autovalores, su rapidez de convergencia disminuye típicamente a medida que nos acercamos a la solución. Alteraciones de tal estructura mediante cambios de variable no ortogonales aceleran la convergencia.

El método clásico de Newton corresponde a $S_k = [F(x_k)]^{-1}$ (inversa del Hessiano), $\alpha_k = 1$. Posee orden dos de convergencia pero su convergencia global no está garantizada y sus requerimientos computacionales, cálculo e inversión del Hessiano a cada iteración, son excesivos. Distintas modificaciones deben ser introducidas para garantizar su convergencia: introducción de un parámetro de búsqueda lineal (de Newton-Raphson) y transformaciones matriciales (Luenberger) para salvar la posible singularidad del Hessiano en puntos lejanos de la solución.

fáciles de aplicar pero solo son indicados en problemas con Hessianos cuasi-diagonales. Su eficacia se altera por rotación de coordenadas pero no por cambios de escala en las variables.

La segunda generación esta constituida por métodos de características intermedias entre el del gradiente y el de Newton. Aceleran la convergencia del primero y disminuyen los requisitos computacionales del segundo. Los hemos clasificados en métodos de direcciones conjugadas y cuasi-Newton según que la modificación se refiera a la dirección de la búsqueda o las aproximaciones a la matriz del Hessiano.

Los distintos métodos de direcciones conjugadas son diferentes extensiones y aplicaciones del método del gradiente conjugado. Desde un punto de vista práctico, lo son en realidad del gradiente conjugado parcial, en el que el proceso se inicializa de nuevo intercalando a intervalos sucesivos una iteración del tipo gradiente clásico. Ello permite conservar la convergencia global y las deseables propiedades de las direcciones conjugadas cerca de la solución. El método de la aproximación cuadrática precisa el cálculo del Hessiano en cada iteración. El de Fletcher-Reeves es el más popular y ha cuasi monopolizado la denominación gradiente conjugado. No requiere el cálculo del Hessiano pero si precisa una búsqueda lineal en cada iteración. El de Polack-Ribière es asimilable al anterior. Todos ellos coinciden cuando se aplican a problemas cuadráticos.

El método PARTAN (tangentes paralelas) se derivó en un principio del llamado gradiente acelerado. Actualmente se considera como un caso particular de direcciones conjugadas,

tico. Converge globalmente y es de fácil aplicación, aunque precisa dos búsquedas lineales por iteración.

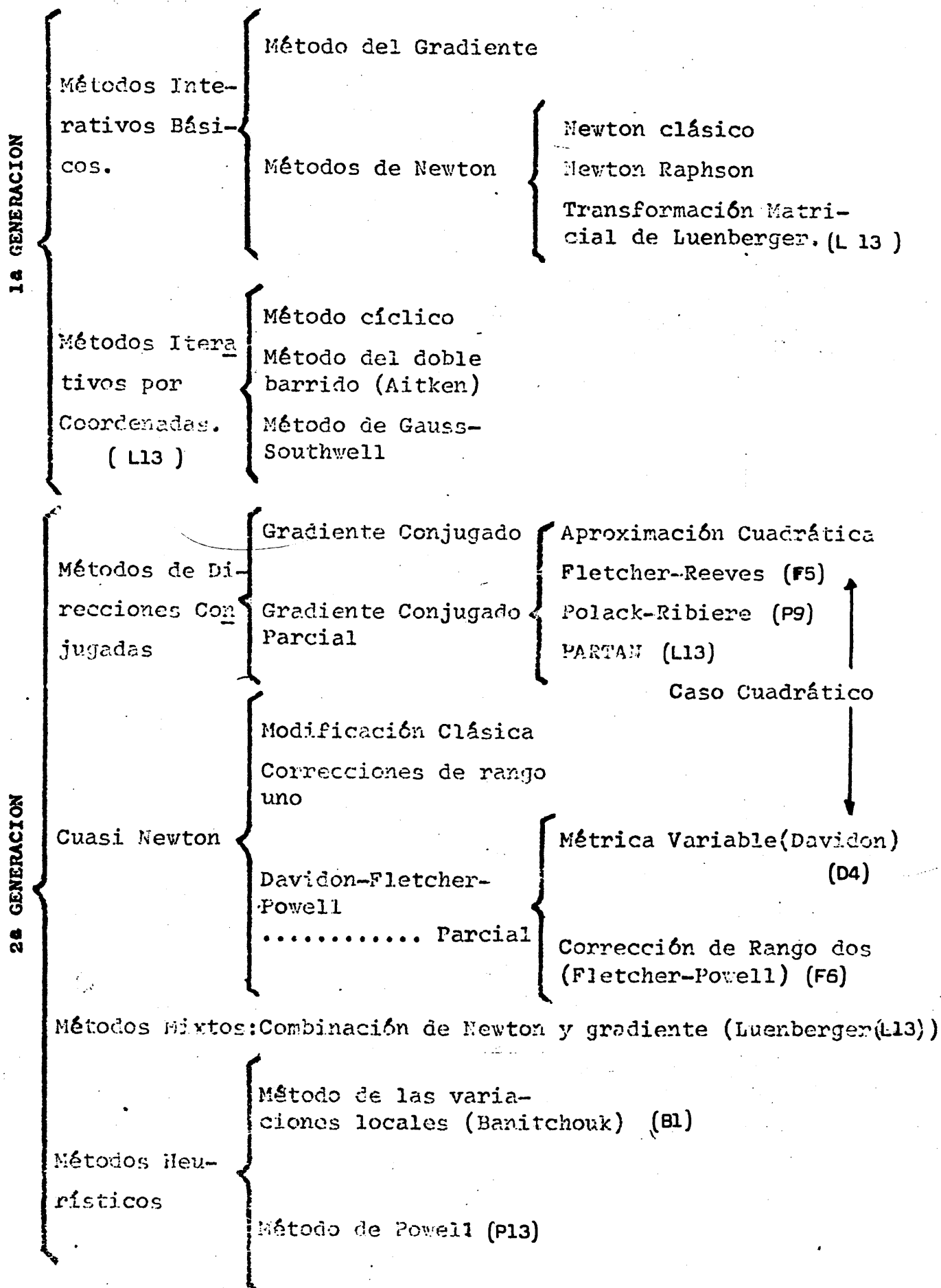
Todos los métodos que llamamos cuasi-Newton usan aproximaciones de la inversa del Hessiano en vez de la verdadera inversa utilizada en los métodos de Newton. El método clásico modificado es el más sencillo y se limita a utilizar el valor inicial de la misma en todas las sucesivas iteraciones. Los métodos más elaborados la construyen usando información almacenada en anteriores iteraciones. El más conocido entre ellos es el de Davidon-Fletcher-Powell (DFP), primeramente propuesto por Davidon bajo el nombre de "métrica variable" y poco después por los otros dos autores. Aunque esta segunda versión ha gozado de mucha mayor difusión, en algunas referencias sigue apareciendo como método de Davidon.

Se le suele llamar también "Corrección de rango dos" porque en cada iteración la inversa del Hessiano es modificada por la suma de dos matrices simétricas de rango uno. El método DFP parcial, incluye reinicializaciones sucesivas. Aplicado al caso cuadrático, coincide con y genera las direcciones del método del gradiente conjugado. Presenta convergencia superlineal. Sus propiedades de convergencia global han sido objeto de animada polémica. Powell presenta una prueba de la misma en el caso convexo.

Los métodos mixtos son combinaciones de los del gradiente y de Newton, empleándose este en los intervalos en los que se conoce el Hessiano.

Los métodos heurísticos reducen o eliminan el uso de de-

mente en la literatura soviética bajo el nombre de variaciones locales. Una completa presentación de los mismos se encuentra en Powell (P13).



Los presentamos clasificados en tres grandes grupos: métodos primales, de modificación y duales, según que en la búsqueda del óptimo analicen el problema original, una aproximación del mismo o su dual.

Los métodos primales son más generalmente aplicables puesto que no presuponen convexidad en el problema. Poseen buenas propiedades de convergencia global, (aunque en sus más simples versiones ninguno de ello sea cerrado) y todos los puntos de la iteración son realizables. En contrapartida exigen una fase previa para localizar una solución inicial realizable y presentan serias dificultades de cálculo. Sus razones de convergencia son particularmente eficientes en el caso de restricciones lineales. Sus componentes principales son los métodos de direcciones realizables, gradiente proyectado y gradiente reducido.

Los métodos de direcciones realizables fueron propuestos por Zoutendijk (22). En sus versiones más simples no son globalmente convergentes debido a los bruscos cambios en el método de generación de las sucesivas direcciones de búsqueda cuando aparece una restricción activa adicional. Los métodos modificados (21, F4) substituyen el concepto de "direcciones realizables" por el de "uniformemente realizables" y hacen intervenir todas las restricciones, activas o no, en los subprogramas lineales que las determinan. Estos establecen un compromiso óptimo entre el deseo de alinear la dirección de búsqueda con el gradiente del objetivo y la necesidad de evitar la frontera del dominio de soluciones posibles. La convergencia global se consigue al coste elevado de resolver en cada iteración un programa

lineal de igual dimensión que el problema original.

La necesidad de reducir estas exigencias computacionales es la base de los otros dos métodos. Ambos tratan de encontrar una dirección de búsqueda realizable que, sin ser óptima localmente, sea más fácil de calcular.

El método del gradiente proyectado (Rosen (R7,)) determina básicamente la dirección de búsqueda en cada iteración proyectando el gradiente sobre el subespacio tangente a la superficie que definen las restricciones activas en ese momento. Aunque es aplicable tanto a restricciones lineales como no lineales , sus características computacionales cambian notablemente en uno o otro caso. En el primero el gradiente proyectado es en sí una dirección realizable y el operador de proyección no precisa ser recalculado completamente en cada iteración. Con restricciones no lineales no siempre es posible alcanzar una nueva solución realizable moviéndose a lo largo del gradiente proyectado, precisándose una nueva proyección sobre la superficie de las restricciones en la dirección perpendicular al plano tangente en el punto original. Ello introduce una serie de dificultades suplementarias, precisa de un proceso iterativo propio y solo puede conseguirse dentro de un cierto grado de aproximación. Por otra parte , ya no es posible actualizar el operador de proyección, sino que debe recalcularse íntegramente en cada iteración. Teóricamente el algoritmo no es globalmente convergente debido a las discontinuidades introducidas en los puntos en que nuevas restricciones pasan a ser activas. Las modificaciones necesarias para prever estas posibles deficiencias destruyen en parte la simplicidad del método.

tendido por Abadie y Carpentier para restricciones no lineales. Se parece al método simplex de programación lineal porque las variables se dividen en básicas y no básicas. El gradiente se llama "reducido" porque en cada iteración se calcula solamente con respecto a las variables no básicas o independientes. Variaciones no lineales de las variables básicas se realizan mediante proyecciones de "retorno" a la superficie definida por las restricciones activas de igual forma que en el gradiente proyectado. Precisa más simples modificaciones (L 13) que este para dotarle de convergencia global y en general requiere menos volumen de cálculo.

Las razones de convergencia de ambos métodos dependen de la estructura de autovalores del Hessiano del Lagrangiano restringido al subespacio tangente a las restricciones activas (Luenberger (L13)).

Una variación del gradiente reducido especialmente usada es el "método simplex convexo" de Zangwill (z1). Su principal característica es que, en cada iteración, solamente una variable independiente varía en la dirección del gradiente reducido.

Los métodos de modificación no son, propiamente hablando, métodos iterativos. El problema original se aproxima por un problema sin restricciones en el que estas se han añadido a la función objetivo. Las dos variantes básicas son los métodos de penalización en los que se especifica un coste elevado por violar las restricciones o los métodos de barrera que favorecen puntos interiores del dominio de soluciones realizables. Estos últimos no son aplicables a restriccio-

Poseen la misma generalidad que los métodos primales y requieren mucha menor complejidad en los programas de cálculo. Sin embargo, a medida que, para mejorar la aproximación, aumentamos el parámetro μ de penalización, disminuye considerablemente la rapidez de convergencia de los algoritmos aplicados. Esencialmente ello es debido a la estructura de autovalores del Hessiano del objetivo modificado. Este posee tantos autovalores que tienden al infinito al hacerlo μ como restricciones activas no degeneradas existan en el punto solución. Esta característica acompaña inevitablemente todo tipo de funciones de penalización y barrera y hace infactible aplicaciones directas del método del gradiente al problema modificado. El método de Newton modificado no se ve afectado en su convergencia por la estructura de autovalores, pero la inversión del Hessiano es todavía mas dificultosa. La aplicación del gradiente conjugado parcial es especialmente efectiva con problemas que poseen un reducido número de restricciones. Asociado con técnicas de normalización (Luenberger (L13)) permiten alcanzar económicamente la razón canónica de convergencia del problema original.

La combinación de métodos de penalización con el gradiente proyectado (Luenberger) permite eliminar los defectos básicos de ambos métodos. El algoritmo resultante es globalmente convergente, su razón de convergencia es la canónica del gradiente proyectado y evita los cálculos del mismo necesarios para mantener la aceptabilidad de la solución.

dimientos iterativos asociándolos con métodos de extrapola-
ción. Estos predicen el punto inicial de la siguiente i-
teración utilizando valores anteriores y sus correspondientes
soluciones.

Bajo la denominación general de métodos duales engloba-
mos a dos categorías disjuntas desde el punto de vista com-
putacional pero en las que la dualidad juega, operativa o
conceptualmente, un importante papel. Solo son aplicables
a aquellos problemas que presentan ciertas propiedades de
convexidad pero compensan esta pérdida de generalidad con
su eficacia en resolver determinados tipos de problemas,
especialmente los de estructura descomponible.

Los métodos de "plano cortante" ("cutting plane") son
creación original de Kelley (K5). Todos ellos consisten
en una serie de programas lineales cuyas soluciones con-
vergen hacia la del problema original. Se diferencian
entre si por el procedimiento utilizado en la construcción
del siguiente programa lineal a partir de la solución ac-
tual. Específicamente, el método de selección del hiper-
plano que separa la actual solución del dominio definido
por las restricciones y la forma en que el poliedro es
actualizado determinan la rapidez de convergencia de los
distintos métodos.

El primero de ellos es el de Kelley, llamado plano cor-
tante convexo, por minimizar una función convexa diferen-
ciable sometida a restricciones de iguales características.
Se parece al método de Newton en que no precisa de ninguna
rutina de búsqueda lineal y coincide con él en problemas

especialmente inadecuado para problemas de elevada dimensionalidad. El método del hiperplano soporte es menos exigente es sus requerimientos de convexidad pero precisa de rutinas de interpolación.

Los métodos duales propiamente dichos resuelven el problema via la solución de su dual. Se basan en la hipótesis de convexidad local de que el Hessiano del Lagrangiano es definido positivo en el punto de solución. En estas condiciones la minimización del problema original es equivalente a la maximización sin restricciones con respecto a las variables duales de la llamada función dual. En la definición de ésta no es preciso incluir las variables duales (multiplicadores) correspondientes a todas las restricciones. Computacionalmente resulta interesante utilizar esta "dualidad parcial" definiendola solamente con respecto a las restricciones igualdad.

En la solución del problema dual pueden ser utilizados los métodos de optimización sin restricciones expuestos. El del gradiente aparece como especialmente conveniente dado que el gradiente de la función dual posee la propiedad de ser inmediatamente calculable una vez esta ha sido obtenida. Pero es preciso tener en cuenta que la evaluación de la función dual implica la resolución de un problema sin restricciones de las variables primales. Ello hace que el método solo sea interesante en los problemas cuya estructura descomponible facilita esta evaluación. (Everett (E1)).

La aplicación de los métodos del tipo gradiente al problema dual tienen su razón de convergencia determinada por

la estructura de autovalores del Hessiano de la función dual en el punto solución. Se demuestra que este es la restricción de la inversa del Hessiano del Lagrangiano en el subespacio determinado por los gradientes de las restricciones en el punto solución. Ello permite definir una razon canónica dual de convergencia en la misma forma que la razon canónica del primal.

La combinación de los métodos duales con las funciones de penalización produce muy potentes algoritmos. Para elevados valores del coeficiente de penalización μ , el Hessiano del dual tiende a $\frac{1}{2\mu} I$ lo cual implica una muy favorable estructura de autovalores en las aplicaciones del método del gradiente y la desaparición de las dificultades computacionales del método de Newton. Asociaciones de este tipo son los algoritmos propuestos por Héstenes (H5) y Powell.

Métodos
Primales

Métodos de Dirección
Realizables.

Método simplificado de Zoutendijk
Métodos generalizados de " "
Método de Topkens y Veinott

Gradiente Proyectado

Restricciones lineales Rosen(R7)
Restricciones no lineales Rosen(R

Gradiente Reducido

Restricciones lineales de Wolfe
Restricciones no lineales de (W2)
Abadie y Carpentier.
Método Simplex Convexo Zangwill(Z1

Métodos
de Modi-
ficación

Métodos de
Penaliza-
ción.

Courant (c9)
Zangwill(Z1)

Método de Newton (L13)
Gradiente Conjugado Normalizado
Luenberger (L13)

Métodos de
Barrera

Carroll
Fiacco-Mc-
Cormick
(F4)

CPEF (Luenberger (L13))
Métodos de Extrapolación (F4)

Métodos del "Plano Cor-
tante"

Plano Cortante Convexo de Kelley
Kelley (K5)

Métodos
Duales

Métodos Duales Propios

Híperplano soporte de Veinott(V5)

Karlin (K2)

Problemas separables Everett(E1)

Combinaciones { Método de los
multiplicadores.
con métodos de Hestenes (H5) .
penalización } Método de Powell

EL PRINCIPIO DE CERTEZA-EQUIVALENCIA EN EL CONTROL OPTIMO
DE SISTEMAS LINEAL-CUADRATICOS ESTOCASTICOS.

Este Apéndice es una presentación resumida del principio de certeza-equivalencia, o teorema de separación, siguiendo la exposición que del mismo hace el profesor A. Bryson en (85) , Capítulo XIV.

Dicho principio se refiere al control óptimo de un sistema cuya dinámica, lineal y afectada por elementos aleatorios, esta representada por las ecuaciones de estado:

$$\dot{X}(t) = F(t)X(t) + G(t)u(t) + w(t) \quad t_0 \leq t \leq T_f$$

donde :

$X(t)$: vector de estado n-dimensional.

$u(t)$: vector de control m-dimensional.

$w(t)$: vector de variables normales puramente aleatorias ("white gaussian noise"), tales que :

$$E[w(t)] = 0$$

$$E[w(t)w'(\tau)] = Q(t)\delta(t-\tau)$$

F, G, Q : Matrices conocidas compuestas por elementos deterministas.

Los estados del sistema no se conocen directamente, sino a través de un vector de medidas $Z(t)$. El proceso de medida , afectado a su vez por elementos aleatorios, esta representado por las ecuaciones :

$$Z(\tau) = H(\tau)X(\tau) + v(\tau) \quad t_0 \leq \tau \leq t$$

donde:

$\mathbf{z}(z)$: vector de mediciones p -dimensional

$\mathbf{H}(z)$: matriz conocida, compuesta por elementos deterministas.

$\mathbf{v}(z)$: vector de variables normales puramente aleatorias ("white gaussian noise "), tales que:

$$E[\mathbf{v}(t)] = 0$$

$$E[\mathbf{v}(t) \mathbf{v}'(\tau)] = \mathbf{R}(t) \delta(t-\tau)$$

Los dos procesos gaussianos puramente aleatorios, $\mathbf{w}(t)$ y $\mathbf{v}(t)$, que afectan respectivamente a la dinámica del sistema y a las medidas efectuadas de su estado, presentan una correlación definida por la matriz \mathbf{T} :

$$E[\mathbf{w}(t) \mathbf{v}'(\tau)] = \mathbf{T}(t) \delta(t-\tau)$$

En conjunto se puede pues escribir:

$$E \begin{bmatrix} \mathbf{w}(t) \\ \mathbf{v}(t) \end{bmatrix} [\mathbf{w}'(\tau), \mathbf{v}'(\tau)] = \begin{bmatrix} \mathbf{Q}(t) & \mathbf{T}(t) \\ \mathbf{T}'(t) & \mathbf{R}(t) \end{bmatrix} \delta(t-\tau)$$

El estado inicial del sistema es un vector de variables aleatorias normales independientes de $\mathbf{v}(t)$, $\mathbf{w}(t)$ de media y matriz de covarianzas conocidas :

$$E[\mathbf{x}(t_0)] = 0$$

$$E[\mathbf{x}(t_0) \mathbf{x}'(t_0)] = \mathbf{P}_0$$

$$E[\mathbf{x}(t_0) \mathbf{v}'(t)] = E[\mathbf{x}(t_0) \mathbf{w}'(t)] = 0$$

El objetivo es la minimización de la funcional cuadrática:

$$J = E \left\{ \frac{1}{2} x'(T_f) S_{T_f} x(T) + \right. \\ \left. + \frac{1}{2} \int_{t_0}^{T_f} [x'(t) u'(t)] \begin{bmatrix} A(t) & N(t) \\ N'(t) & B(t) \end{bmatrix} \begin{bmatrix} x(t) \\ u(t) \end{bmatrix} dt \right\}$$

donde $A(t)$, $B(t)$, S_{T_f} , $N(t)$ son matrices conocidas y $A(t)$ es semidefinida positiva y $B(t)$ positiva definida.

Se trata de determinar el control óptimo $u(t)$ a aplicar al sistema en cada instante t , no como una función del estado del sistema en dicho instante $x(t)$, puesto que este es imposible de conocer exactamente, sino como una función de todas las mediciones $z(\tau)$ efectuadas de dicho estado en todos los instantes τ comprendidos entre el inicial t_0 y el actual t : $t_0 \leq \tau \leq t$

Tal control óptimo viene dado por la relación :

$$u(t) = - C(t) \hat{x}(t)$$

donde :

$$C(t) = B^{-1}(t) [G'(t) S(t) + N'(t)]$$

$\hat{x}(t)$ se obtiene de la integración del sistema de ecuaciones diferenciales :

$$\dot{\hat{x}}(t) = F(t) \hat{x}(t) + K(t) [z(t) - H(t) \hat{x}(t)]$$

con las condiciones en los límites :

$$\hat{x}(t_0) = 0$$

y la matriz $S(t)$ obedece al sistema de ecuaciones diferen-

ciales con las correspondientes condiciones en los límites:

$$\dot{S} = -SF - F'S + C'BC - A$$

$$S(t_f) = S_f$$

y la matriz $K(t)$ se obtiene en cada instante mediante la relación:

$$K(t) = [P(t)H'(t) + T(t)]R^{-1}(t)$$

en la que la matriz $P(t)$ obedece al sistema de ecuaciones diferenciales :

$$\dot{P} = FP + PF' - K RK' + Q$$

con las condiciones en los límites :

$$P(t_0) = P_0$$

Como se observa en las anteriores ecuaciones, el control óptimo en el caso estocástico así definido, obedece a la misma forma estructural que el control óptimo en el caso determinista, substituyendo los valores reales del vector de estado, (ahora desconocidos), por las estimaciones de los mismos representadas por el vector $\hat{X}(t)$. El cálculo de los mismos (filtro de Kalman) se efectúa, a partir de los valores $Z(\tau), (t_0 \leq \tau \leq t)$ mediante las anteriores ecuaciones.

REFERENCIAS

ABREVIACIONES UTILIZADAS EN LAS REFERENCIAS

AER:	American Economic Review
AT:	Automatika y Telemekhanika
CI:	Computer Journal
ECSR:	Economic Cybernetics Studies and Research
IEE:	Institute of Electric Engineers
EJ:	Economic Journal
IEEE-TAC:	Institución of Electrical and Electronical Engineers
IER:	International Economic Review
IJC:	International Journal on Control
IJSS:	International Journal in System Sciences
JET:	Journal of Economic Theory
JOTA:	Journal of Optimization Theory and Applications
JPE:	Journal of Political Economy
QJE:	Quarterly Journal of Economics
OR:	Operations Research
RES:	Review of Economic Studies
RFEP:	Revue Francaise d'Economie Politique
RIRO:	Revue d'Informatique et de Recherche Operationnelle
PJM:	Pacific Journal of Mathematics
SIAM-JC:	Society for Industrial and Applied Mathematics Journal on Control
WEJ:	Wertern Economic Journal
ZVMMF:	Zh. vychisl matem. i matem. FIZIKI

- A2 Abadie J. , Carpentier J.
"GENERALIZATION OF THE WOLFE REDUCED GRADIENT METHOD TO THE CASE OF NON LINEAR CONSTRAINTS"
en "OPTIMIZATION"; Fletcher R.
Academic Press 1969
- A3 Abadie J.
"APPLICATION OF THE GRG ALGORITHM TO OPTIMAL CONTROL PROBLEMS"
en "INTEGER AND NON LINEAR PROGRAMMING"
North Holland 1970
- A4 Abadie J. , Bichara M.
"EXPERIENCES NUMERIQUES SUR UNE METHODE DE RESOLUTION DE CERTAINS PROBLEMES DE COMMANDE OPTIMALE"
Electricité de France Nota HI 739/00 1972
- A5 Albouy M. , Breton A.
"INTERPRETATION ECONOMIQUE DU PRINCIPE DU MAXIMUM"
RIRO 14, 1968 p 37-68
- A6 Albouy M.
"MODELES DE CROISSANCE ET TAUX D'ACTUALISATION"
RFEP Marzo, Abril 1970 p 260-286
- A7 Albouy M.
"LA REGULATION ECONOMIQUE DANS L'ENTREPRISE"
Dunod 1972
- A8 Armand L.
"LA DESCENTELIZATION PAR LES PRIX"
METRA Septiembre 1968
- A9 Armand L.
"INTERPRETATION ECONOMIQUE DU PRINCIPE DU MAXIMUM DE PONTRYAGIN"
METRA 3, 1969
- A10 Athams M. , Kendrick D.
"ECONOMICS AND CONTROL: A SURVEY"
IEEE-TAC Octubre 1974 p 518

- "NUMERICAL SOLUTIONS OF PROBLEMS WITH VARIATIONAL LIMITS BY THE METHOD OF LOCAL VARIATIONS"
ZVMMF 6(2), 1968 p. 149-173.
- B2 Bellman R.
"DYNAMIC PROGRAMMING"
Princeton University Press 1957

"LA PROGRAMMATION DYNAMIQUE ET SES APPLICATIONS"
Dunod 1965.
- B3 Benavie A.
"THE ECONOMICS OF THE MAXIMUM PRINCIPLE"
WEJ Vol. II, Diciembre 1975.
- B4 Boudarel R. , Delmas J. , Guichet P.
"COMMANDE OPTIMALE DES PROCESSUS"
Dunod 1969.
- B5 Bryson A. , Ho Y.C.
"APPLIED OPTIMAL CONTROL"
Ginn Editor 1969.
- B6 Brody A.
"PROPORTIONS, PRICES AND PLANNING"
North Holland 1972.
- B7 Brody A.
"OPTIMAL AND TIME OPTIMAL PATH OF THE ECONOMY"
en "CONTRIBUTIONS TO INPUT/OUTPUT ANALYSIS"
Carter y Brody
North Holland 1970.
- B8 Brolley T.
"A DYNAMIC MODEL FOR THE BELGIUM ECONOMY"
ECONOMICS OF PLANNING Vol. 9 nº1,2 1969
p. 155.
- B9 Burmeister, Dobell.
"TEORIA MATEMATICA DEL CRECIMIENTO ECONOMICO"
Editorial Bösch, Barcelona.
- C1 Canon, Cullum, Polack.
"OPTIMAL CONTROL AND MATHEMATICAL PROGRAMMING"
Mc Graw Hill 1970.

- C3 Case J.H.
"TOWARDS A THEORY OF MANY PLAYERS DIFERENTIAL GAMES"
IJC Vol. 7, 1969 p. 179-197.
- C4 Cass D.
"OPTIMUM GROWTH IN AN AGREGATIVE MODEL OF CAPITAL ACCUMULATION: A TURNPIKE THEOREM"
ECONOMETRICA Octubre 1966.
- C5 Chakravarty S.
"CAPITAL AND DEVELOPMENT PLANNING"
M.I.T. Press 1969.
- C6 Chow C.G.
"PROBLEMS OF ECONOMIC POLICY FROM THE VIEW-POINT OF OPTIMAL CONTROL"
Research Memorandum nº 139
Econometric Research Program Princeton University Mayo 1972.
- C7 Chow C.G.
"OPTIMAL CONTROL OF LINEAR ECONOMETRIC SYSTEMS WITH FINITE TIME HORIZON"
IER Vol. 13 nº1 Febero 1972 p. 16-25.
- C8 Citron S.J.
"ELEMENTS OF OPTIMAL CONTROL"
Holt & Rinehart Inc. 1969.
- C9 Courant R.
"CALCULUS OF VARIATIONS AND SUPPLEMENTARY NOTES"
New-York University 1962.
- C10 Courtin P.
"PERFORMANCE INDEX SENSITIVITY IN OPTIMAL CONTROL PROBLEMS"
IEEE TAC XVI Junio 1971.
- C11 Cullum J.
"DISCRETE APPROXIMATIONS TO CONTINUOUS OPTIMAL CONTROL PROBLEMS"
SIAM-JC 7, 1969 p. 32-49.

- D1 Dacuna R.O. , Polack E.
"CONSTRAINED MINIMIZATION UNDER VECTOR-
VALUED CRITERIA IN FINITE-DIMENSIONAL
SPACES"
Universidad de California, Berkeley
Memorandum ERL-M188 1966.
- D2 Daniel J.W.
"ON THE CONVERGENCE OF A NUMERICAL METHOD
FOR OPTIMAL CONTROL PROBLEM"
JOTA 4, 1969 p. 330-342
"THE APPROXIMATE MINIMIZATION OF FUNCTIONALS"
Prentice Hall 1971.
- D3 Dantzig G.B.
"LINEAR PROGRAMMING AND EXTENSIONS"
Princeton University Press 1963.
- D4 Davidon W.C.
"VARIABLE METRIC METHODS FOR MINIMIZATION"
AEC Research and Development Report 5990
1959.
- D5 Denn M.M.
"OPTIMIZATION BY VARIATIONAL METHODS"
Mc Graw Hill 1969.
- D6 Dreyfus S.
"DYNAMIC PROGRAMMING AND THE CALCULUS OF
VARIATIONS"
Academic Press 1964.
- D7 Dixit A.
"OPTIMAL DEVELOPMENT IN THE LABOR SURPLUS
ECONOMY"
RES 35, 1969 p. 23-34.
- D8 Dobell R.
"SOME CHARACTERISTIC FEATURES OF OPTIMAL
CONTROL PROBLEMS IN ECONOMIC THEORY"
IEEE TAC Vol. 14 nº1 Febrero 1969.
- D9 Dobell R. , Ho Y.C.
"OPTIMAL INVESTMENT POLICY: AN EXAMPLE OF
A CONTROL PROBLEM IN ECONOMIC THEORY"
IEEE TAC Vol. 12 Febrero 1967.
- D10 Dobell R.
"OPTIMIZATION IN MODELS OF ECONOMIC GROWTH"
en "STUDIES IN OPTIMIZATION"
SIAM Vol. 1, 1970 p. 1-27.

"LINEAR PROGRAMMING AND ECONOMIC ANALYSIS"
Mc Graw Hill, 1958.

D12 Dorfman R.

"AN ECONOMIC INTERPRETATION OF OPTIMAL
CONTROL THEORY"
AER Diciembre 1969 p. 817-831.

D13 Dubobsiskii S.V.

"CONSTRUCTION OF AN OPTIMAL ECONOMIC PLAN"
AT nº 8 Agosto 1972 p. 101-114.

D14 Dubobsiskii S.V., Uzdemir A.P.

"OPTIMALITY CRITERIA AND VARIATIONAL AP-
PROACHES IN DYNAMIC MODELS OF AN ECONOMY"
AT, nº 6, Junio 1974 p. 90-98.

D15 Dyukalov A.N., Ivanov Y.N., Tokarev V.V.

"CONTROL THEORY FOR ECONOMICS SYSTEMS" I,II
AT I nº 5, Mayo 1974 p. 117-132
II nº 6, Junio 1974 p. 69-89.

E1 Everett H.

"GENERALIZED LAGRANGE MULTIPLIER METHOD FOR
SOLVING PROBLEMS OF OPTIMUM ALLOCATION OF
RESOURCES"
OR 11, 1963, p. 399-413.

F1 Faure P., Huard P.

"RESOLUTION DES PROGRAMMES MATHEMATIQUES A
FONCTION NON LINEAIRE PAR LA METHODE DU
GRADIENT REDUIT"
RIRO 9, 1965 p. 167-205.

F2 Fedorenko N.

"ECONOMICO-MATHEMATICAL MODELS IN SOVIET-
ECONOMIC PLANNING"
MATEKON 1969 p. 337-352.

F3 Fedoseev A.V.

"OPTIMAL CONTROL IN THE THREE SECTORS MODEL
OF AN UNDERDEVELOPED COUNTRY"
ZVMEF 12,4, 1972 p. 925-942.

F4 Fiacco A.V. , Mc Cormack G.P.

"NON LINEAR PROGRAMMING"
John Wiley 1968.

"FUNCTION MINIMIZATION BY CONJUGATE GRADIENTS"

CJ 7(2), 1964. p. 149-154.

F6 Fletcher R. , Powell M.J.

"A RAPIDLY CONVERGENT DESCENT METHOD FOR MINIMIZATION"

CJ. 6, 1963 p. 163-168.

F7 Fox, Sengupta, Thorbecke

"THE THEORY OF QUANTITATIVE ECONOMIC POLICY"
North Holland 1966.

F8 Friedman B.M.

"OPTIMAL ECONOMIC STABILIZATION POLICY: AN EXTENDED FRAMEWORK"

JPE 80, 1972 p. 1002-1021.

G1 Gelfand I. , Fomin S.

"CALCULUS OF VARIATIONS"

Prentice Hall

New-York 1963.

G2 Geoffrion A.M.

"PRIMAL-RESOURCE-DIRECTIVE APPROACHES FOR OPTIMIZING NON LINEAR DECOMPOSABLE SYSTEMS"

Working Paper 141

The Rand Corporation 1968.

G3 Goodwin R.M.

"THE OPTIMAL GROWTH PATH FOR AN UNDERDEVELOPED ECONOMY"

EJ. Diciembre 1961.

G4 Grateloup G. , Titli A.

"A COMBINED DECOMPOSITION AND COORDINATION METHOD IN LARGE DIMENSION OPTIMIZATION PROBLEMS"

IJSS 1973 Vol. 4 n° 4, p. 577-595.

G5 Grateloup G. , Titli A.

"TWO-LEVEL DYNAMIC OPTIMIZATION METHODS"

JOTA Vol. 15 n° 5 1975 p. 533-547.

G6 Gunckel-Franklin

"A GENERAL SOLUTION FOR LINEAR SAMPLED-DATA CONTROL SYSTEM"

Transaction ASME Vol. 85D, 1963 p. 197.

"OPTIMAL CONTROL FOR SYSTEMS DESCRIBED BY
DIFFERENCE EQUATIONS"
Collection Control and Dynamic Systems, Vol. 1,
Academic Press 1964.

H2 Helmer J.Y.

"LA COMMANDE OPTIMALE EN ECONOMIE"
Dunod 1972.

H3 Hestenes M.R.

"OPTIMIZATION THEORY"
John Wiley 1975.

H4 Hestenes M.R.

"CALCULUS OF VARIATIONS AND OPTIMAL CON-
TROL THEORY"
John Wiley 1966.

H5 Hestenes M.R.

"MULTIPLIER AND GRADIENT METHODS"
JOTA Vol. 4 n°5 1969 p. 303-320.

H6 Ho Y.C.

"DIFFERENTIAL GAMES, DYNAMIC OPTIMIZATION
AND GENERALIZED CONTROL THEORY"
JOTA Vol. 6 n° 3 1970 p. 179-207.

H7 Holt C.

"LINEAR DECISION RULES FOR ECONOMIC STABI-
LIZATION AND GROWTH"
QJE Vol. 76 1962 p. 20-45.

H8 Holtzman J.M.

"CONVEXITY AND THE MAXIMUM PRINCIPLE FOR
DISCRETE SYSTEMS"
IEEE TAC Enero 1966.

I1 Institut de Recherche en Informatique Appliquée

"METHODES NUMERIQUES D'ANALYSE DE SYSTEMES"
Tomo I
IRIA Cahier n° 9 1972.

I2 Isaacs R.

"DIFFERENTIAL GAMES"
John Wiley 1965.

- J2 Jordan, Polack
"OPTIMAL CONTROL OF APERIODIC DISCRETE-TIME SYSTEMS"
SIAM J.C. Vol. 2 1964.
- J3 Jorgenson D.W.
"THE STRUCTURE OF MULTI-SECTOR DYNAMIC MODELS"
IER Vol. 2 1961 p. 276-293.
- K1 Kamien M. , Schwartz N.
"SUFFICIENT CONDITIONS IN OPTIMAL CONTROL THEORY"
JET Vol. 3:2 Junio 1971 p. 207-214.
- K2 Karlin
"MATHEMATICAL METHODS AND THEORY IN GAMES, PROGRAMMING AND ECONOMICS"
Vol. I
Addison-Wesley 1959.
- K3 Kaufman A. , Cruon R.
"LA PROGRAMMATION DYNAMIQUE"
Dunod 1965.
- K4 Katsenelinboigen A.I.
"OPTIMAL CONTROL AND THE PRICE MECHANISM"
MATEKON Vol. III, nº 3 1970.
- K5 Kelley
"THE CUTTING-PLANE METHOD FOR SOLVING CONVEX PROGRAMS"
SIAM Journal VIII, 4, 1960 p. 703-712.
- K6 Kendrick D.A. , Taylor L.J.
"A DYNAMIC NON-LINEAR PLANNING MODEL FOR KOREA"
en "PRACTICAL APPROACHES TO DEVELOPMENT PLANNING"
I. Adelman
The John Hopkins Press 1969.

K8 Kendrick D.A.

"MULTIPERIOD MODELS FOR NATIONAL PLANNING"
Texas University 1974.

K9 Kirillova F.M.

"ON THE CORRECTNESS OF THE FORMULATION OF
AN OPTIMAL CONTROL PROBLEM"
SIAM JC 1, 1963 p. 224-239.

K10 Kirk D.E.

"OPTIMAL CONTROL THEORY. AN INTRODUCTION"
Prentice Hall 1970.

K11 Kornai, Liptack

"MATHEMATICAL PROGRAMMING AND STRUCTURAL
DECISIONS"
North Holland 1967.

K12 Kuhn, Szegö

"DIFFERENTIAL GAMES AND RELATED TOPICS"
North Holland 1971.

K13 Kurtz, M.

"THE GENERAL INSTABILITY OF A CLASS OF
GROWTH PROCESSES"
RES Abril 1968 p. 155-179.

K14 Kurtz, M.

"OPTIMUM PATH OF CAPITAL ACCUMULATION UN-
DER THE MINIMUM TIME OBJECTIVE"
ECONOMETRICA 1965.

L1 Larson

"STATE INCREMENT DYNAMIC PROGRAMMING"
American Elsevier 1970.

L2 Lasdon L.S. , Fox R.L. , Ratner M.W.

"NON-LINEAR OPTIMIZATION USING THE GENERA-
LIZED REDUCED GRADIENT METHOD"
RIRO 8, Noviembre 1974 p. 73-104.

"DUALITY AND DECOMPOSITION IN MATHEMATICAL PROGRAMMING"

IEEE Transactions on Systems, Man and Cybernetics 4, nº 2 1968 p. 86-100.

L4 Lasdon L.S.

"DESIGN AND TESTING OF A GRG CODE FOR NON-LINEAR OPTIMIZATION"

Report TM-353 Case Western University
Marzo 1975.

L5 Lasdon L.S.

"OPTIMIZATION THEORY FOR LARGE SYSTEMS"

Mc Millan 1970.

L6 Leitmann G. , Liu P.T.

"A DIFFERENTIAL GAME MODEL OF LABOR-MANAGEMENT NEGOTIATION DURING A STRIKE"

JOTA Vol. 13 nº 4 1974 p. 427-435.

L7 Livesey D.A.

"THE SINGULARITY PROBLEM IN THE DYNAMIC INPUT/OUTPUT MODEL"

IISS Abril 1973.

L8 Livesey D.A.

"CAN MACRO-ECONOMIC PLANNING PROBLEMS EVER BE TREATED AS A QUADRATIC REGULATOR PROBLEM?"

IFAC/IFORS International Conference
IEE Publication Number 101.

L9 Livesey D.A.

"THE IMPORTANCE OF NUMERICAL ALGORITHMS FOR SOLVING ECONOMIC OPTIMIZATION PROBLEMS"

IJSS 1974 Vol. 5, nº 5 p. 435-451.

L10 Livesey D.A.

"CONTROL THEORY AND INPUT/OUTPUT ANALYSIS"

IJSS 1971.

L11 Luenberger D.

"MATHEMATICAL PROGRAMMING AND CONTROL THEORY: TRENDS OF INTERPLAY"

en Geoffrion "PERSPECTIVES IN OPTIMIZATION"
Addison-Wesley 1972.

L12 Luenberger D.

"OPTIMIZATION BY VECTOR SPACES METHODS"

John Wiley 1969.

"INTRODUCTION TO LINEAR AND NON-LINEAR PROGRAMMING"
Addison-Wesley 1973.

M1 Mc Fadden D.

"ON THE EXISTENCE OF OPTIMAL DEVELOPMENT
PROGRAMS IN INFINITE HORIZON ECONOMIES"
"Proceedings" de la IEA Conferencia
Mc Millan 1973.

M2 Mangasarian

"SUFFICIENT CONDITIONS FOR THE OPTIMAL CONTROL OF NON-LINEAR SYSTEMS"
SIAM JC Vol. 4 1966.

M3 Manne A.

"MULTI VECTOR MODELS FOR DEVELOPMENT PLANNING"
Frontiers in Quantitative Economics Tomo 2
North Holland 1974.

M4 Martes, Pyndick

"AN APPLICATION OF OPTIMAL CONTROL TO INVESTMENT ALLOCATION FOR DEVELOPMENT PLANNING"
International Institute for Quantitative Economics
Montreal 1972.

M5 Mehra R.K. , Davis R.E.

"A GENERALIZED GRADIENT METHOD FOR OPTIMAL CONTROL PROBLEM WITH INEQUALITY CONSTRAINTS AND SINGULAR ARCS"
IEEE TAC Vol. AC-17 nº1 1972 p. 69-79.

M6 Mera K.

"AN EMPIRICAL DETERMINATION OF A DYNAMIC UTILITY FUNCTION"
RES Vol. 50 nº 1 1968.

M7 Mesarovic, Macko, Takahara

"THEORY OF HIERARCHICAL MULTILEVEL SYSTEMS"
Academic Press 1970.

M8 Miele A. , Hoesley, Levy, Coggins

"ON THE METHOD OF MULTIPLIERS FOR MATHEMATICAL PROGRAMMING PROBLEMS"
JOTA Vol. 10 nº 1 1972 p. 1-30.

"APPLICATION OF THE OPTIMAL CONTROL THEORY
TO ECONOMIC ANALYSIS"
ECSR Vol. 3, 1973 p.99.

M10 Morishima

"TEORIA DEL CRECIMIENTO ECONOMICO"
Editorial Tecnos.

M11 Murakami Y.

"EFFICIENT PATHS OF ACCUMULATION AND THE
TURNPIKE OF THE JAPANESE ECONOMY"
en "APPLICATIONS OF INPUT/OUTPUT ANALYSIS"
Carter y Brody capítulo 2
North Holland 1970.

N1 National Bureau of Economic Research (NBER)

"ANNALS OF ECONOMIC AND SOCIAL MEASUREMENT"
Vol. 3 nº 1 1974
Vol. 1 nº 4 1972

N2 Newberry D.M.

"PUBLIC POLICY IN THE DUAL ECONOMY"
E.J. 82, 1972 p. 567-589.

N3 Norman A., Norman M.

"THE COMPUTATION OF DETERMINISTIC OPTIMAL
MACROECONOMIC POLICY"
1973 Winter Meeting of the Econometric So-
ciety.

P1 Pallu de la Barriere

"COURS D'AUTOMATIQUE THEORIQUE"
Dunod 1968.

P2 Pant P.

"NOTES ON PERSPECTIVE OF DEVELOPMENT OF
INDIA: 1960-61 to 1975-76"
Planning Commission
New Delhi 1964.

P3 Pau L.F.

"DIFFERENTIAL GAMES AMONG SECTORS IN A MA-
CROECONOMY"
IFAC/IFORS International Conference
Warwick Julio 1973
IEE PublicationNumber 101.

"DYNAMIC DECOMPOSITION TECHNIQUES"
en Wisner
Mc Graw Hill 1971.

P5 Peterson D.W.

"ON SENSIBILITY IN OPTIMAL CONTROL PROBLEMS"
JOTA Vol. 13, nº 1 1974
"THE ECONOMIC SIGNIFICANCE OF AUXILIARY
FUNCTIONS IN OPTIMAL CONTROL"
IER Vol. 14, nº 1 Febrero 1973.

P6 Phillips A.W.

"MATHEMATICAL MODELS IN ECONOMIC DYNAMICS"
ECONOMICA Vol. 17 1950 p. 382-395.

P7 Phillips A.W.

"STABILIZATION POLICY IN A CLOSED ECONOMY"
E.J. nº 64, 1954 p. 290.

P8 Phillips A.W.

"STABILIZATION POLICY AND THE TIME FORM OF
LAGGED RESPONSES"
E.J. nº 67 1957 p. 265.

P9 Polack E.

"COMPUTATIONAL METHODS IN OPTIMIZATION"
Academic Press 1971.

P11 Polack E.

"AN HISTORICAL SURVEY OF COMPUTATIONAL ME-
THODS IN OPTIMAL CONTROL"
SIAM Review Vol. 15 nº 2 Abril 1973 .
p. 553-584.

P12 Pontryagin, Boltyanskii, Gankrelidze, Mishchenko

"THE MATHEMATICAL THEORY OF OPTIMAL PROCES-
SES"
John Wiley 1962.

P13 Powell M.J.

"A SURVEY OF NUMERICAL METHODS FOR UNCONS-
TRAINED OPTIMIZATION"
SIAM Review 12 (1), 1970 (79,97)

P14 Pyndick

"OPTIMAL PLANNING FOR ECONOMIC STABILIZA-
TION"
North Holland 1973.

"OPTIMAL GROWTH IN A LINEAR-LOGARITHMIC
ECONOMY"

IER 1960 p. 1-33.

R2 Ramsey R.P.

"A MATHEMATICAL THEORY OF SAVINGS"
Economic Journal XXXVIII, 152
Vol. 38 Diciembre 1928 p. 543.

R3 Rockafellar R.T.

"SADDLE POINTS OF HAMILTONIAN SYSTEMS IN
CONVEX PROBLEMS OF LAGRANGE"
JOTA 1973.

R4 Rockafellar R.T.

"GENERALIZED HAMILTONIAN EQUATIONS FOR CON-
VEX PROBLEMS OF LAGRANGE"
PJM 33,2 1970 p. 411-427.

R5 Rockafellar R.T.

"DUALITY AND STABILITY IN EXTREMUM PROBLEMS
INVOLVING CONVEX FUNCTIONS"
PJM 21, 1967 p. 167-187.

R6 Rockafellar R.T.

"THE MULTIPLIER METHOD OF HESTENES AND PO-
WELL APPLIED TO CONVEX PROGRAMMING"
JOTA Vol. 12 nº 6 1973 p. 555-562.

R7 Rosen J.B.

"THE GRADIENT PROJECTION METHOD FOR NON-
LINEAR PROGRAMMING"
SIAM Journal 8 (1) 1960 p. 181-217.

R8 Rosen J.B.

"OPTIMAL CONTROL AND CONVEX PROGRAMMING"
en "NON-LINEAR PROGRAMMING" Abadie
North Holland 1967 p. 289-302.

S1 Schoeffler J.D.

"STATIC MULTILEVEL SYSTEMS"
en Wisner capítulo 1, obra citada.

S2 Shell K.

"ESSAYS ON THE THEORY OF OPTIMAL ECONOMIC
GROWTH"
M.I.T. Press 1967.

"ON THE SOLUTION OF THE OPEN-LOOP NASH RICATTI EQUATIONS IN LINEAR-QUADRATIC DIFFERENTIAL GAMES"

IJC 1973 Vol. 18 n° 1 p. 57-63.

S4 Simon H.A.

"DYNAMIC PROGRAMMING UNDER UNCERTAINTY WITH A QUADRATIC CRITERION FUNCTION"
ECONOMETRICA Vol. 24 1956 p. 74-81.

S5 Smirnov A.D.

"PROBLEMS OF CONSTRUCTING AN OPTIMAL INTER-BRANCH MODEL OF SOCIALIST REPRODUCTION"
en "CONTRIBUTIONS TO INPUT/OUTPUT ANALYSIS"
Carter y Brody editores
North Holland 1970.

S6 Solow R.

"GROWTH THEORY: AN EXPOSITION"
Oxford Clarendon Press 1970.

S7 Starr A.W., Ho Y.C.

"NON ZERO-SUM DIFFERENTIAL GAMES"
JOTA Vol. 3 n° 3 1969 p. 185-219.

S8 Stern N.H.

"OPTIMUM DEVELOPMENT IN A DUAL ECONOMY"
RES 1972 p. 171-183.

S9 Stoléru L.G.

"A QUANTITATIVE MODEL OF GROWTH OF THE ALGERIAN ECONOMY"
PhD. Thesis Stanford University 1963.

S10 Stoléru L.G.

"AN OPTIMAL POLICY FOR ECONOMIC GROWTH"
ECONOMETRICA 1965.

S11 Stone J.R.N.

"TRANSITIONAL PLANNING: THE ADAPTATION OF THE ECONOMY TO A HIGHER RATE OF GROWTH"
en "ON POLITICAL ECONOMY AND ECONOMETRICS... ESSAYS IN HONOUR OF OSKAR LANGE"
Warsaw 1965.

T1 Tabak D. , Kuo B.C.

"OPTIMAL CONTROL BY MATHEMATICAL PROGRAMMING"
Prentice Hall 1971.

"MATHEMATICAL ECONOMICS"
Dryden Press 1974.

T3 Tamura H.

"APPLICATIONS OF DUALITY AND DECOMPOSITION
IN HIGH ORDER MULTI-STAGE DECISION PROBLEMS"
Cambridge University Engineering Department
Report n° CVED/B-Control TR 49 1973.

T4 Terceiro J.

"ERROR ANALYSIS OF THE LINEAR STOCHASTIC
CONTROL PROBLEM"
Proceedings de la 8th Hawaiian Conference
on System Science p. 66-69
Western Periodicals Company Enero 1975.

T5 Theil H.

"A NOTE ON CERTAINTY EQUIVALENCE IN DYNAMIC
PLANNING"
ECONOMETRICA Vol. 25 1957 p. 346-349.

T6 Theil H.

"OPTIMAL DECISION RULES FOR GOVERNMENT AND
INDUSTRY"
North Holland 1964.

T7 Tou S.I.

"ON LINEAR CONTROL THEORY"
Transaction AIEE parte III
Vol. 80 n° 18 1961.

T8 Tsukui J.

"APPLICATION OF A TURNPIKE THEOREM TO PLAN-
NING FOR EFFICIENT ACCUMULATION. AN EXAM-
PLE FOR JAPAN"
ECONOMETRICA Enero 1968.

T9 Turnovsky S.

"OPTIMAL STABILIZATION POLICIES FOR DETER-
MINISTIC AND STOCHASTIC LINEAR ECONOMIC
SYSTEMS"
R.E.S. 1973, 40, 79.

T10 Tustin A.

"THE MECHANISM OF ECONOMIC SYSTEMS"
James Heinemann Inc. 1953.

"THE PROBLEM OF LAGRANGE WITH DIFFERENTIAL
INEQUALITIES AS ADDED SIDE CONDITIONS"
University of Chicago Press 1937.

V2 Varaiya R.P.

"ON THE EXISTENCE OF SOLUTION TO A DIFFE-
RENTIAL GAME"
IJC Vol. 5 1967 p. 153-162.

V3 Vegara J.

"PROGRAMACION MATEMATICA Y CALCULO ECONOMICO"
Ed. Vicens Vives 1976.

V4 Veinott A.F.

"THE SUPPORTING HYPERPLANE METHOD FOR UNI-
MODAL PROGRAMMING"
OR XV, 1, 1967 p. 147-152.

V5 Von Weizacker C.C.

"EXISTENCE OF OPTIMAL PROGRAMS OF ACCUMULA-
TION FOR AN INFINITE TIME HORIZON"
RES Vol. 32 nº 2 Abril 1965 p. 85-104.

W1 Wang C. , Fan Liang

"THE DISCRETE MAXIMUM PRINCIPLE"
John Wiley 1964.

W2 Wismer D.A.

"OPTIMIZATION METHODS FOR LARGE SCALE SYS-
TEMS... WITH APPLICATIONS"
Mc Graw Hill 1971.

W3 Wolfe P.

"METHODS OF NON-LINEAR PROGRAMMING"
en "NON-LINEAR PROGRAMMING" Abadie, editor
North Holland 1967.

W4 Wonham W.M.

"STOCHASTIC PROBLEMS IN OPTIMAL CONTROL"
IEEE International Convention Record
4a parte 1963
"ON THE SEPARATION THEOREM OF STOCHASTIC
CONTROL"
SIAM JC Vol. 6, nº 2 Mayo 1968.

"NON-LINEAR PROGRAMMING: A UNIFIED APPROACH"
Prentice Hall 1969.

Z2 Zoutendijk G.

"METHODS OF FEASIBLE DIRECTIONS"
Elsevier 1960.
